8-1-2011

Directional Subdifferentials and Optimality Conditions

Ivan Ginchev
Technical University of Varna, Bulgaria, vechnig@yahoo.com

Boris S. Mordukhovich
Wayne State University, boris@math.wayne.edu

DIRECTIONAL SUBDIFFERENTIALS AND OPTIMALITY CONDITIONS

IVAN GINCHEV and BORIS S. MORDUKHOVICH

WAYNE STATE UNIVERSITY

Detroit, MI 48202

Department of Mathematics
Research Report

2011 Series
#8

This research was partly supported by the USA National Science Foundation

Directional subdifferentials and optimality conditions

Ivan Ginchev* Boris S. Mordukhovich†

Abstract

This paper is devoted to the introduction and development of new dual-space constructions of generalized differentiation in variational analysis, which combine certain features of subdifferentials for nonsmooth functions (resp. normal cones to sets) and directional derivatives (resp. tangents). We derive some basic properties of these constructions and apply them to optimality conditions in problems of unconstrained and constrained optimization.

Key words: variational analysis, directional normals and subgradients, calculus rules, nonsmooth optimization, necessary and sufficient optimality conditions.

Mathematics Subject Classification: 49J52, 49J53, 90C29.

1 Introduction

Variational analysis and generalized differentiation is largely developed via a geometric dual-space approach in the book by Mordukhovich [11] with numerous applications collected in the second volume [12]. This approach has some advantages with respect to other constructions. For instance, the basic subdifferential of [11] is of smaller size in comparison with other useful constructions with developed calculus rules, which makes it appropriate for a larger set of various mathematical and applied problems, particulary those in optimization. At the same time, the basic subdifferential of [11] and the other constructions associated defined directly in dual spaces are nonconvex-valued and hence cannot be generated via duality by any primal-space construction (like directional derivatives and tangent cones). Observe however that primal-space constructions involving tangency and directions play an important role in some aspects of variational and nonsmooth analysis, especially in finite-dimensional spaces; see, e.g., [15] and the references therein. Therefore, preserving the dual space approach of [11] on one hand and wishing to have the capability to treat within this approach tangency and directions, it is desirable to introduce directionally dependent dual notions. It is the primary goal of this paper, which develops some ideas and results preliminary announced in [7].

*University of Insubria, Department of Economics, Via Monte Generoso 71, 21100 Varese, Italy, and Technical University Varna, Department of Mathematics, 9010 Varna, Bulgaria (vechnig@yahoo.com). Research of this author was partly supported by a grant of Technical University Varna.

†Wayne State University, Department of Mathematics, Detroit, MI 48202, USA (boris@math.wayne.edu) and King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia. Research of this author was partly supported by the USA National Science Foundation under grant DMS-1007132, by the Australian Research Council under grant DP-1292508, and by the Portuguese Foundation of Science and Technologies under grant MAT/11109.

In what follows we define directional normals to nonempty sets and directional subgradients of general nonsmooth functions on Banach spaces based on dual-space constructions while involving directions in the primal space. To the best of our knowledge, the idea of a directional subdifferential (collections of subgradients) related to Clarke's generalized gradient of Lipschitzian functions was first explored by Chaney [2, 3] primarily in the framework of second-order constructions and applications to second-order optimality conditions . It was further developed on some abstract level in [5] and applied, in particular, to derive the second-order optimality conditions from [1] and [4]. Recently the notions of directional subdifferentials appear in [13]. In all these papers the primal notions (directional derivatives, tangents) are essential parts of the definitions.

This paper defines directional normals and subgradients following another scheme with the omission of primal constructions. Extending the results from [7], we derive necessary and sufficient optimality conditions via the new directional subdifferential and normal cone. These conditions have their counterparts in terms of directional derivatives; e.g., in [6] with Dini derivatives, and in [4] and [9] with Hadamard derivatives. Among the advantages of the new directional dual-space constructions, which differ them from the "nondirectional" counterparts [11], we particularly mention their remarkable behavior under scalarization, close relationships with strict directional derivatives of a new type, special properties for certain classes of functions depending on directions, etc. Probably the most important results of the paper show that the usage of the directional subdifferentials and normals allows us to derive necessary optimality conditions in problems of unconstrained and constrained optimization that can be more efficient than those in [12] and other publications. Furthermore, we are able to establish some sufficient optimality conditions in terms the new directional constructions, which is not the case for their nondirectional analogs.

The rest of the paper is organized as follows. In Section 2 we present some background material and define directional normals to sets; in fact those with respect to a certain set of directions. Then we pass in Section 3 to our basic directional subdifferential construction and establish some of its remarkable properties mentioned above. Section 4 contains a number of calculus rules for the directional subdifferential and its specifications applied to favorable classes of functions. In Section 5 we study unconstrained optimization problems in finite-dimensional and infinite-dimensional spaces, derive necessary and sufficient optimality conditions in terms of our basic directional subdifferential and its modifications, and present several examples illustrating their applications and comparison with known results. Finally, Section 6 deals with problems of constrained optimization with general inequality constraints and establishes new necessary optimality conditions via the directional subdifferential. We also present several motivating and illustrating examples, which shed light on future developments.

2 Directional Normals to Sets

In this paper we basically use the standards notation of variational analysis; see, e.g., [11, 15]. Unless otherwise stated, all the spaces are real Banach normed by $\|\cdot\|$. Usually with X we denote such a space. The open balls in X with center x_0 and radius r are $B(x_0,r) := \{x \in X \mid \|x-x_0\| < r\}$; the unit ball centered at the origin is $B := B(0,1)$. The corresponding closed balls are $\overline{B}(x_0,r) := \{x \in X \mid \|x-x_0\| \le r\}$ and $\overline{B} := \overline{B}(0,1)$. The topological dual of X is X^*, and $\langle\cdot,\cdot\rangle$ is the canonical pairing on $X^* \times X$. The weak* topology on X^* is denoted by w^*. In what follows Ω stands for a nonempty subset of X.

Given a set-valued mapping $F\colon X \rightrightarrows Y$ between Banach spaces (we write as usual $F\colon X \to Y$ when F is single-valued) and given some point $x_0 \in \operatorname{dom} F := \{x \in X|\ F(x) \neq \emptyset\}$, the *Painlevé-Kuratowski upper/outer limit* of F at x_0 is defined by

$$\operatorname*{Lim\,sup}_{x \to x_0} F(x) := \big\{y \in Y\big|\ \exists \text{ sequences } x_k \to x_0,\ y_k \to y \text{ with } y_k \in F(x_k) \text{ as } k \in \mathbb{N}\big\}. \tag{2.1}$$

It is important to emphasize here that, when $F\colon X \rightrightarrows X^*$ is a mapping between a Banach space X and its topological dual, the convergence of $y_k = x_k^* \in X^*$ to $y = x^*$ in (2.1) is always *sequential* in the *weak** topology of X^*, and we denote it by $x_k^* \xrightarrow{w^*} x^*$. Given $\Omega \subset X$ and $x_0 \in \Omega$, write $x \xrightarrow{\Omega} x_0$ if $x \to x_0$ with $x \in \Omega$. Using this notation, we recall (see [11, Definition 1.1]) that

$$\widehat{N}_\varepsilon(x_0;\Omega) := \Big\{x^* \in X^*\Big|\ \limsup_{x \xrightarrow{\Omega} x_0} \frac{\langle x^*, x - x_0\rangle}{\|x - x_0\|} \le \varepsilon\Big\}, \quad \varepsilon \ge 0, \tag{2.2}$$

is the collection of ε-*normals* to Ω at $x_0 \in \Omega$. For $\varepsilon = 0$ in (2.2), the set $\widehat{N}(x_0;\Omega) := \widehat{N}_0(x_0;\Omega)$ is known as the *Fréchet/regular normal cone* (or prenormal cone) to Ω at x_0. Observe that the ε-normal set (2.2) is convex for any $\varepsilon \ge 0$; however, it does not possesses satisfactory calculus rules even for simple nonconvex sets Ω in finite dimensions. Furthermore, the regular normal cone $\widehat{N}(x_0;\Omega)$ may often be trivial ($=\{0\}$) as, e.g., in the case of $\widetilde{\Omega} := \{(x_1,x_2) \in \mathbb{R}^2 \mid x_2 \ge -|x_1|\}$ at $x_0 = (0,0)$. This does not look natural for a normal cone notion.

The situation changes dramatically by employing a sequential regularization of the the set-valued mapping $F(x,\varepsilon) := \widehat{N}_\varepsilon(x;\Omega)$ from X to X^* via (2.1). The collection of normals

$$N(x_0;\Omega) := \operatorname*{Lim\,sup}_{x \xrightarrow{\Omega} x_0,\, \varepsilon \to 0^+} \widehat{N}_\varepsilon(x;\Omega) \tag{2.3}$$

obtained in this way is known as the *Mordukhovich basic/limiting normal cone* to Ω at $x_0 \in \Omega$. If the space X is Asplund (i.e., each if its separable subspace has a separable dual; for example, every reflexive space) and if the set Ω is locally closed around x_0 (i.e., its intersection with some closed ball centered at x_0 is closed), then we can equivalently put $\varepsilon = 0$ in (2.3); see [11, Theorem 2.25]. Furthermore, in the case of finite-dimensional spaces we have the equivalent representation

$$N(x_0;\Omega) = \operatorname*{Lim\,sup}_{x \to x_0} \big[\operatorname{cone}\big(x - \Pi(x;\Omega)\big)\big] \tag{2.4}$$

via the Euclidean projector $\Pi(x;\Omega) := \{w \in \Omega \mid \|x - w\| = \operatorname{dist}(x;\Omega)\}$, where the symbol "cone" stands for the (generally nonconvex) cone spanned on the set in question, and where $\operatorname{dist}(\cdot;\Omega)$ denotes the usual distance function of a set. Note that (2.4) was actually the original definition of the normal cone in [10]. It is easy to observe that the limiting normal cone (2.3) can be *nonconvex* in simple finite-dimensional settings; e.g., for the set $\widetilde{\Omega}$ presented above. This means that it *cannot* be dual/polar to any tangential approximation of a set, since polarity always generates convexity. On the other hand, this normal cone and subdifferential/coderivative constructions for functions and set-valued mappings associated with it enjoy comprehensive calculus rules based on the *variational/extremal principles* of variational analysis; see [11] for more details.

We intend to implement the idea of using *directions* in the primal space X to improve the basic dual-space construction (2.3). Since directional approximations of sets in the primal space are naturally formalized via tangents to sets at given points, let us recall and make it used a powerful tangential approximation of sets known as the *Bouligand-Severi tangent/contingent cone* to Ω at x_0, which is defined by

$$T(x_0;\Omega) := \operatorname*{Lim\,sup}_{t\to 0+} \frac{\Omega - x_0}{t} \tag{2.5}$$

via the outer limit (2.1) taken with respect to the norm topology of X; see [11, Definition 1.8].

To proceed further, take a subset $Q \subset X$ and consider the conic hull

$$C_Q := \operatorname{cone} Q := \big\{\lambda q \in X \big|\ \lambda \geq 0,\ q \in Q\big\}$$

generated by the set Q. Abbreviating the notation, put $C_u = C_{\{u\}}$. We write $x \overset{\Omega,Q}{\to} x_0$ if $x \overset{\Omega}{\to} x_0$ and simultaneously $\operatorname{dist}(\frac{x-x_0}{\|x-x_0\|}; C_Q) \to 0$ with the convention that $\frac{x-x_0}{\|x-x_0\|} = 0$ for $x = x_0$. When $Q = \{u\}$ for some $u \in X \setminus \{0\}$, write $x \overset{\Omega,u}{\to} x_0$ instead of $x \overset{\Omega,\{u\}}{\to} x_0$. Observe that $x \overset{\Omega,Q}{\to} x_0$ is equivalent to $x \overset{\Omega,C_A}{\to} x_0$, a consequence of $C_{C_Q} = C_Q$. Also $x \overset{\Omega,X}{\to} x_0$ is the same as $x \overset{\Omega}{\to} x_0$.

Now we define directional counterparts of generalized normals from (2.2) and (2.3).

Definition 2.1 (generalized normals with respect to sets). *Let $x_0 \in \Omega \subset X$, and let $Q \subset X$. For $\delta > 0$ put $Q_\delta := Q + \delta B$. Then:*

(i) *Given $\varepsilon \geq 0$, the collection of* ε-NORMALS WITH RESPECT TO Q *to the set Ω at $x_0 \in \Omega$ is*

$$\widehat{N}_{\varepsilon,Q}(x_0;\Omega) := \widehat{N}_\varepsilon\big(x_0;\Omega \cap (x_0 + C_Q)\big) = \Big\{x^* \in X^* \Big|\ \limsup_{x \overset{\Omega\cap(x_0+C_Q)}{\to} x_0} \frac{\langle x^*, x - x_0\rangle}{\|x - x_0\|} \leq \varepsilon\Big\}. \tag{2.6}$$

(ii) *The* LIMITING NORMAL CONE WITH RESPECT TO Q *to the set Ω at $x_0 \in \Omega$ is*

$$N_Q(x_0;\Omega) := \operatorname*{Lim\,sup}_{x \overset{\Omega,Q}{\to} x_0,\, \varepsilon\to 0+,\, \delta \to 0+} \widehat{N}_{\varepsilon,Q_\delta}(x;\Omega). \tag{2.7}$$

If $Q = \{u\}$ in (2.6) and (2.7), then corresponding constructions are referred as to normals in DIRECTION *$u \in X$ to Ω at x_0; cf.* [7, Definition 2.3].

It follows from (2.6), (2.7), and (2.5) that $\widehat{N}_{\varepsilon,Q}(x_0;\Omega) = X^*$ if $\Omega \cap (x_0 + C_Q) = \{x_0\}$ and that $N_u(x_0;\Omega) = X^*$ if $u \notin T(x_0;\Omega)$. Furthermore, for $Q = X$ the constructions from Definition 2.1 reduce to (2.6) and (2.7), respectively, i.e., $\widehat{N}_{\varepsilon,X}(x_0;\Omega) = \widehat{N}_\varepsilon(x_0;\Omega)$ and $N_X(x_0;\Omega) = N(x_0;\Omega)$. As in the proof of [11, Theorem 2.25], we can equivalently put $\varepsilon = 0$ in (2.7) if X is Asplund and the set Ω is locally closed around x_0.

It is possible to check, following the proofs in [11], that the limiting normal cone with respect to sets (2.7) largely possesses properties similar to those for the basic normal cone (2.3). We are not going to discuss them here, since our focus in this paper is mainly on directional subdifferentials and optimality conditions in their terms considered below.

3 Directional Subdifferentials of Functions

In this section, given a subset $\Omega \subset X$ of a Banach space and a function $\varphi\colon \Omega \to \mathbb{R}$, we construct a directional subdifferential of φ at points $x_0 \in \Omega$. Recall that the Mordukhovich basic/limiting subdifferential of φ at x_0 is defined by (see [11, Definition 1.77])

$$\partial\varphi(x_0) := \{x^* \in X^* \,|\, (x^*, -1) \in N((x_0, \varphi(x_0)); \operatorname{epi}\varphi)\} \tag{3.1}$$

via the basic normal cone (2.3) to the epigraph

$$\operatorname{epi}\varphi = \{(x, y) \in X \times \mathbb{R} \,|\, x \in \Omega,\, y \geq \varphi(x)\}.$$

Now we proceed similarly to define the corresponding directional subdifferential with replacing the basic normal cone in (3.1) by its directional counterpart (2.7).

Definition 3.1 (basic directional subdifferential). *Consider a function $\varphi\colon \Omega \to \mathbb{R}$ on a set $\Omega \subset X$ and a point $x_0 \in \Omega$. Taking a direction $0 \neq u \in X$ and abbreviating $\{u\} \times \mathbb{R}$ as $u \times \mathbb{R}$, define the* (*basic*) DIRECTIONAL SUBDIFFERENTIAL *of φ at x_0 in direction u by*

$$\partial_u\varphi(x_0) := \{x^* \in X^* \,|\, (x^*, -1) \in N_{u\times\mathbb{R}}((x_0, \varphi(x_0)); \operatorname{epi}\varphi)\}. \tag{3.2}$$

We easily get from definitions (3.2) and (2.5) that $\partial_u\varphi(x_0) = X^*$ if $u \notin T(x_0, \Omega)$. Observe also from the definitions that every directional subgradient $x^* \in \partial_u\varphi(x_0)$ can be described as follows: there are sequences $x_k \overset{\Omega,u}{\to} x_0$, $\varepsilon_k \to 0^+$, $\delta_k \to 0^+$, and $x_k^* \overset{w^*}{\to} x^*$ with

$$x_k^* \in \widehat{N}_{\varepsilon_k,(u+\delta_k B)\times\mathbb{R}}((x_k, \varphi(x_k)); \operatorname{epi}\varphi) = \widehat{N}_{\varepsilon_k}((x_k, \varphi(x_k)); \operatorname{epi}\varphi \cap ((x_k, \varphi(x_k)) + C_{(u+\delta_k B)} \times \mathbb{R})).$$

The latter means in turn that for any number $\varepsilon > 0$ there exists $\delta(k) > 0$ such that whenever $x \in D_u(x_k, \delta(k), \delta_k) := B(x_k, \delta(k)) \cap (x_k + C_{u+\delta_k B})$ we have

$$\langle x_k^*, x - x_k\rangle - (r - \varphi(x_k)) \leq (\varepsilon_k + \varepsilon)(\|x - x_k\| + |r - \varphi(x_k)|) \ \text{ for all } \ r \geq \varphi(x). \tag{3.3}$$

Similarly to (3.3) in the sequel we use the notation

$$D_u(x, \Delta_0, \Delta_1) := B(x, \Delta_0) \cap (x + C_{u+\Delta_1 B}) \ \text{ with the closure } \ \overline{D}_u(x, \Delta_0, \Delta_1). \tag{3.4}$$

It is worth mentioning that our approach allows us to introduce the subdifferential of a function at a point with respect to a set, which is defined via the corresponding normal cone (2.7).

Definition 3.2 (subdifferential with respect to sets). *Consider a function $\varphi\colon \Omega \to \mathbb{R}$ with $\Omega \subset X$, a point $x_0 \in X$, and a set $Q \subset X$. Then*

$$\partial_Q\varphi(x_0) := \{x^* \in X^* \,|\, (x^*, -1) \in N_{Q\times\mathbb{R}}((x, \varphi(x)); \operatorname{epi}\varphi)\} \tag{3.5}$$

is the LIMITING SUBDIFFERENTIAL WITH RESPECT TO Q *of φ at x_0.*

We can see that construction (3.5) unifies both the Mordukhovich basic/limiting subdifferential (3.1) corresponding to $\partial\varphi(x_0) = \partial_X\varphi(x_0)$ and the basic directional subdifferential (3.2) corresponding to $\partial_u\varphi(x_0) = \partial_{\{u\}}\varphi(x_0)$. Some of the properties of our basic directional subdifferential established below can be generalized to the subdifferential with respect to sets, although in the this paper we pay the main attention to the theory and applications of the (basic) directional subdifferential (3.2) and some of its modifications used in deriving optimality conditions.

First observe the following useful properties of the directional subdifferential used in the sequel. Pick $0 \neq u \in X$, $0 < \delta < \|u\|$, and $x - x_0 \in C_{u+\delta B}$. Then $x - x_0 = \lambda(u + \delta b)$ with some $b \in B$ and $\lambda \geq 0$. We have the estimates $\|u\| - \delta \leq \|u + \delta b\| \leq \|u\| + \delta$, and hence

$$\lambda = \frac{\|x - x_0\|}{\|u + \delta b\|} = \frac{\|x - x_0\|}{(\|u\| + \theta)}, \quad -\delta \leq \theta \leq \delta. \tag{3.6}$$

Consequently for $x \neq x_0$ we have the relationships

$$\frac{x - x_0}{\|x - x_0\|} - \frac{u}{\|u\|} = \left(\frac{1}{\|u\| + \theta} - \frac{1}{\|u\|}\right) u + \frac{\delta}{\|u\| + \theta} b \to 0 \quad \text{as} \quad \delta \to 0^+. \tag{3.7}$$

The latter allows us to describe behavior of the directional subdifferential with respect to an appropriate subgradient scalarization.

Theorem 3.3 (directional subdifferential under scalarization). *Let $x^* \in \partial_u \varphi(x_0)$, and let $y^* \in X^*$ be such that $\langle y^*, u\rangle \leq \langle x^*, u\rangle$. Then we have $y^* \in \partial_u \varphi(x_0)$.*

Proof. Take $\varepsilon_k \to 0^+$, $\delta_k \to 0^+$, $x_k \overset{\Omega,u}{\to} x_0$, and $x_k^* \overset{w^*}{\to} x_0$ such that for each $\varepsilon > 0$ there is $\delta(k) > 0$ for which (3.3) is satisfied with ε substituted by $\varepsilon/2$ and all $x \in D_u(x_0, \delta(k), \delta_k)$. Put $y_k^* := x_k^* + y^* - x^* \overset{w^*}{\to} y^*$. Then $x_k^* = y_k^* + x^* - y^*$ and

$$\langle y_k^*, x - x_k\rangle + \langle x^* - y^*, x - x_k\rangle - \big(r - \varphi(x_k)\big) \leq \left(\varepsilon_k + \frac{\varepsilon}{2}\right)\Big(\|x - x_k\| + |r - \varphi(x_k)|\Big)$$

whenever $r \geq \varphi(x)$. The above inequality yields (3.3) with x_k^* substituted by y_k^* provided that $\langle y^* - x^*, x - x_k\rangle \leq \frac{1}{2}\varepsilon\|x - x_k\|$. The latter obviously holds when $x = x_k$. For $x \neq x_k$ it is true for sufficiently large k since

$$\begin{aligned}\langle y^* - x^*, x - x_k\rangle &= \left\langle y^* - x^*, \frac{x - x_k}{\|x - x_k\|} - \frac{u}{\|u\|}\right\rangle \|x - x_k\| + \frac{\|x - x_k\|}{\|u\|}\langle y^* - x^*, u\rangle \\ &\leq \|y^* - x^*\| \cdot \left\|\frac{x - x_k}{\|x - x_k\|} - \frac{u}{\|u\|}\right\| \cdot \|x - x_0\|\end{aligned}$$

and $\frac{x - x_k}{\|x - x_k\|} - \frac{u}{\|u\|} \to 0$ as $x \to x_k$, which is a consequence of $x - x_k \in C_{u+\delta_k B}$ and $\delta_k \to 0^+$; see (3.7). This completes the proof of the theorem. $\square$

The next useful result ensures the closedness of the directional subdifferential (3.2) for arbitrary functions on finite-dimensional spaces.

Proposition 3.4 (closed values of the directional subdifferential). *Let $\varphi\colon \Omega \to \mathbb{R}$, let $x_0 \in \Omega \subset X$, and let $u \in X \setminus \{0\}$. Assume that $\dim X < \infty$. Then the set $\partial_u \varphi(x_0)$ is closed in X.*

Proof. For $u \notin T(x_0, \Omega)$ we have $\partial_u \varphi(x_0) = X$, and the closedness conclusion is obvious. Suppose now that $u \in T(x_0, \Omega)$ and consider sequences $x_\nu^* \to x^*$ as $\nu \to \infty$ with $x_\nu^* \in \partial_u \varphi(x_0)$ for all $\nu \in \mathbb{N}$. It follows from (3.3) that for each fixed ν there are sequences $\varepsilon_{k\nu} \to 0^+$, $\delta_{k\nu} \to 0^+$, $x_{k\nu} \overset{\Omega,u}{\to} x_0$, and $x_{k\nu}^* \to x_\nu^*$ such that whenever $\varepsilon > 0$ we can find $\delta(k, \nu) > 0$ with the property

$$\langle x_{k\nu}^*, x - x_{k\nu}\rangle - \big(r - \varphi(x_{k\nu})\big) \leq (\varepsilon_{k\nu} + \varepsilon)\big(\|x - x_{k\nu}\| + |r - \varphi(x_{k\nu})|\big) \tag{3.8}$$

for all $r \geq \varphi(x)$, $x \in D_u(x_{k\nu}, \delta(k,\nu), \delta_{k\nu})$, and large $k \in \mathbb{N}$. Employing further the diagonal process allows us to select a subsequence ν_k along which

$$\varepsilon_{k\nu_k} \to 0^+,\ \delta_{k\nu_k} \to 0^+,\ x_{k\nu_k} \overset{\Omega,u}{\to} x_0, \text{ and } x^*_{k\nu_k} \to x^*$$

as $k \to \infty$. Hence relationship (3.8) holds whenever $\nu = \nu_k$ and $x \in D_u(x_{k\nu_k}, \delta(k,\nu_k), \delta_{k\nu_k})$ for all $k \in \mathbb{N}$ is sufficiently large. Employing again the directional subgradient description (3.3), we conclude that $x^* \in \partial_u\varphi(x_0)$ and thus complete the proof of the proposition. $\square$

Note that the above closedness property of Proposition 3.4 does not generally hold in infinite dimensions; it can be demonstrated similarly to [11, Example 1.7] given in the case of the basic normal cone (2.3) (or subdifferential of the indicator function) in Hilbert spaces. The proof of Proposition 3.4 is violated in infinite dimensions due to the fact that the diagonal process can not be applied to the weak* topology of X^*.

It turns out that the directional subdifferential defined in the dual space X^* could be connected to an appropriate directional derivative construction defined in the primal space X. To proceed, we introduce two directional derivative versions used in what follows.

Definition 3.5 (strict directional derivatives). *Let $\varphi\colon \Omega \to \mathbb{R}$, let $x_0 \in \Omega \subset X$, and let $u \in T(x_0;\Omega)$. Then the* UPPER STRICT DIRECTIONAL DERIVATIVE $\varphi'_+(x_0;u)$ *and the* STRICT DIRECTIONAL DERIVATIVE $\varphi'(x_0;u)$ *of φ at x_0 in direction u are defined, respectively, by*

$$\varphi'_+(x_0;u) := \limsup_{x \overset{\Omega,u}{\to} x_0,(t,v)\to(0^+,u)} \frac{1}{t}\Big(\varphi(x+tv) - \varphi(x)\Big),$$
$$\varphi'(x_0;u) := \lim_{x \overset{\Omega,u}{\to} x_0,(t,v)\to(0^+,u)} \frac{1}{t}\Big(\varphi(x+tv) - \varphi(x)\Big)$$

where the pair $(t,v) \in \mathbb{R}_+ \times X$ is chosen so that $x + tv \in \Omega$. For convenience we put $\varphi'_+(x_0;u) := \varphi'(x_0,u) = \infty$ if $u \notin T(x_0;\Omega)$.

First we establish the following relationship between the directional subdifferential and the upper strict directional derivative.

Theorem 3.6 (directional subdifferential and and upper strict directional derivative). *Given $\varphi\colon \Omega \to \mathbb{R}$, $\Omega \subset X$, and $x_0 \in \Omega$, we have*

$$\sup\langle \partial_u\varphi(x_0), u\rangle \leq \varphi'_+(x_0;u) \ \textit{for any}\ u \in X. \tag{3.9}$$

Proof. If $\varphi'_+(x_0;u) = \infty$, inequality (3.9) is trivially satisfied. Assume now that $\varphi'_+(x_0;u)$ is finite. Take $x^* \in \partial_u\varphi(x_0)$ and put $x := x_k + tv$ in the subgradient description (3.3), where $0 < t < \delta(k)/(\|u\| + \delta_k)$ and $v \in u + \delta_k B$ are such that $x = x_k + tv \in \Omega$. It is easy to check that $x \in D_u(x_k, \delta(k), \delta_k)$, while for $r := \varphi(x) = \varphi(x + tv)$ it follows from (3.9) that

$$\langle x^*_k, v\rangle \leq \frac{1}{t}\Big(\varphi(x_k + tv) - \varphi(x_k)\Big) + (\varepsilon_k + \varepsilon)\Big(\|v\| + \frac{1}{t}|\varphi(x_k + tv) - \varphi(x_k)|\Big). \tag{3.10}$$

Taking 'limsup' with respect to $k \to \infty$ and $(t,v) \to (0^+,u)$, we get $\langle x^*, u\rangle \leq \varphi'_+(x_0;u) + \varepsilon(\|u\| + |\varphi'_+(x_0;u)|)$ provided that

$$\lim_{k\to\infty} \langle x^*_k, v\rangle = \langle x^*, u\rangle. \tag{3.11}$$

Since $\varepsilon > 0$ is arbitrary and since $\varphi'_+(x_0;u)$ is finite, it follows from the above that $\langle x^*, u\rangle \leq \varphi'_+(x_0;u)$, which justifies the desired estimate (3.9) in the case when $\varphi'_+(x_0;u)$ is finite. To complete the proof in this case, we check now that the limiting relationship (3.11) is satisfied. Observe that $\langle x_k^*, w\rangle \to \langle x^*, w\rangle$ as $k \to \infty$ for all $w \in X$ by the weak* convergence and that the set $\{\|x_k^*\|\}_k$ is bounded in X^* due the Banach-Steinhaus uniform boundedness principle. Thus

$$|\langle x_k^*, v\rangle - \langle x^*, u\rangle| \leq \|x_k^*\| \cdot \|v-u\| + |\langle x_k^*, u\rangle - \langle x^*, u\rangle| \to 0 \;\text{ as }\; k \to \infty.$$

which justifies (3.11) and hence (3.9) in the case of $|\varphi'_+(x_0;u)| < \infty$.

It remains to consider the case of $\varphi'_+(x_0;u) = -\infty$ and hence $u \in T(x_0;\Omega)$. Let us check that $\partial_u\varphi(x_0) = \emptyset$. Arguing by contradiction, suppose the opposite, pick an arbitrary element $x^* \in \partial_u\varphi(x_0;u)$, and choose a sequence $x_k^* \xrightarrow{w^*} x^* \in \partial_u\varphi(x_0)$ as above. Passing to the limsup in (3.10) with $\varepsilon < 1$ gives us $-\infty < \langle x^*, u\rangle \leq -\infty$, a contradiction. Remembering the usual convention that the supremum over an empty set is $-\infty$, we justify formula (3.9) in the last case and thus complete the proof of the theorem. $\square$

The next observation provides an expected precise relationship between the directional subdifferential and strict directional derivative for the case of linear continuous functionals.

Proposition 3.7 (directional subdifferential and strict directional derivative of linear functions). *Consider a function $\varphi : X \to \mathbb{R}$ given by $\varphi(x) := \langle \xi, x\rangle$, where $\xi \in X^*$ is fixed. Then $\xi \in \partial_u\varphi(x_0)$ whenever $x_0 \in X$ and $0 \neq u \in X$, and we have the representation*

$$\partial_u\varphi(x_0) = \big\{x^* \in X^* \big|\, \langle x^*, u\rangle \leq \varphi'(x_0;u)\big\} \;\text{ with }\; \varphi'(x_0;u) = \langle \xi, u\rangle. \tag{3.12}$$

Proof. The formula for the strict directional derivative $\varphi'(x_0,u) = \langle \xi, u\rangle$ easily follows from

$$\frac{1}{t}(\varphi(x+tv) - \varphi(x)) = \langle \xi, v\rangle.$$

To show further that $\xi \in \partial_u\varphi(x_0)$, take $x_k^* = \xi$ and get from (3.3) that

$$-(r - \langle \xi, x\rangle) \leq (\varepsilon_k + \varepsilon)(\|x - x_k\| + |r - \langle \xi, x_k\rangle|) \;\text{ for all }\; r \geq \langle \xi, x\rangle.$$

The inclusion "$\subset$" in (3.12) follows now from Theorem 3.6. To prove the opposite inclusion "$\supset$", let $x^* \in X^*$ with $\langle x^*, u\rangle \leq \langle \xi, u\rangle$ and show that $x^* \in \partial_u\varphi(x_0)$. Consider some sequences $\varepsilon_k \to 0^+$, $x_k \xrightarrow{\Omega,u} x_0$, and $x_k^* = x^* \xrightarrow{w^*} x^*$. Picking $\varepsilon > 0$, we wish to show that whenever $\delta_k \to 0^+$ and $\delta(k) > 0$ representation (3.3) holds for all $x \in D_u(x_k, \delta(k), \delta_k)$; it is written now as

$$\langle x^* - \xi, x - x_k\rangle - (r - \langle \xi, x\rangle) \leq (\varepsilon_k + \varepsilon)(\|x - x_k\| + |r - \langle \xi, x_k\rangle|) \;\text{ for all }\; r \geq \langle \xi, x\rangle.$$

The latter inequality is obvious for $x = x_k$. For $x \neq x_k$ it can be rewritten as

$$\Big\langle x^* - \xi, \frac{x - x_k}{\|x - x_k\|}\Big\rangle - \frac{r - \langle \xi, x\rangle}{\|x - x_k\|} \leq (\varepsilon_k + \varepsilon)\Big(1 + \frac{|r - \langle \xi, x_k\rangle|}{\|x - x_k\|}\Big) \;\text{ whenever }\; r \geq \langle \xi, x\rangle,$$

which holds for all $k \in \mathbb{N}$ sufficiently large since

$$\begin{aligned}\Big\langle x^* - \xi, \frac{x - x_k}{\|x - x_k\|}\Big\rangle &= \Big\langle x^* - \xi, \frac{x - x_k}{\|x - x_k\|} - \frac{u}{\|u\|}\Big\rangle + \Big\langle x^* - \xi, \frac{u}{\|u\|}\Big\rangle \\ &\to \frac{1}{\|u\|}\Big(\langle x^*, u\rangle - \langle \xi, u\rangle\Big) \leq 0 \;\text{ as }\; k \to \infty.\end{aligned}$$

This completes the proof of the proposition. $\square$

4 Calculus of Directional Subgradients

In this section we derive some calculus rules for the new directional subdifferential (3.2) used in what follows. More results can be obtained in this way similarly to the calculus of the basic subdifferential (3.1); cf. [11]. Roughly speaking, a major difference between constructions (3.1) and (3.2) is that the latter mainly concerns *directional* properties and the corresponding classes of functions, e.g., directional strict differentiable, directional Lipschitzian, etc.

We start with the following simple while important observation.

Proposition 4.1 (multiplication by nonnegative constants). *Let $\varphi\colon \Omega \to \mathbb{R}$, $x_0 \in \Omega \subset X$, and $0 \neq u \in X$. Then we have*

$$\partial_u(\lambda\varphi)(x_0) = \lambda\partial_u(\varphi)(x_0) \;\; \textit{for any} \;\; \lambda > 0. \tag{4.1}$$

Proof. It is sufficient to justify the inclusion "$\supset$ in (4.1), which yields the converse one being applied to $\varphi(x) = (1/\lambda)\varphi(\lambda x)$. To proceed, pick any $x^* \in \partial_u\varphi(x_0)$ and $\varepsilon > 0$. By (3.2) select sequences $\varepsilon_k \to 0^+$, $\delta_k \to 0^+$, $x_k \overset{\Omega,u}{\to} x_0$, and $x_k^* \overset{w^*}{\to} x^*$ as $k \to \infty$ and find $\delta(k) > 0$ such that inequality (3.3) holds for all $x \in D_u(x_k, \delta(k), \delta_k)$ with ε replaced by by $\varepsilon/\max(1,\lambda)$. Multiplying the latter inequality by λ and replacing λr by r, we get

$$\begin{aligned}\langle\lambda x_k^*, x - x_k\rangle - (r - \lambda\varphi(x_k) &\leq \Big(\varepsilon_k + \frac{\varepsilon}{\max(1,\lambda)}\Big)\big(\lambda\|x - x_k\| + |r - \lambda\varphi(x_k)|\big)\\ &\leq \big(\max(1,\lambda)\varepsilon_k + \varepsilon\big)\big(\|x - x_k\| + |r - \lambda\varphi(x_k)|\big) \;\; \text{for all} \;\; r \geq \lambda\varphi(x).\end{aligned}$$

Since $\lambda x_k^* \overset{w^*}{\to} \lambda x^*$ as $k \to \infty$, the displayed formula shows that $\lambda x^* \in \partial_u(\lambda\varphi)(x_0)$, which thus completes the proof of the proposition. $\square$

To proceed further, recall that a real function φ defined in a neighborhood of $x_0 \in X$ is *strictly differentiable* at x_0 if there exists an element $\varphi'(x_0) \in X^*$, the strict derivative of f at x_0, such that whenever $\varepsilon > 0$ there is $\delta > 0$ with the property

$$|\varphi(y) - \varphi(x) - \varphi'(x_0)(y - x)| \leq \varepsilon\|y - x\| \;\; \text{for all} \;\; x, y \in B(x_0, \delta).$$

The next definition extends the latter property introducing its directional variant.

Definition 4.2 (strictly directionally differentiable functions). *Let $\varphi\colon \operatorname{dom}\varphi \to \mathbb{R}$ with $x_0 \in \operatorname{dom}\varphi \subset X$ on a Banach space X, and let the set $D_u(x, \Delta_0, \Delta_1)$ be defined in (3.4). Then φ is* STRICTLY DIRECTIONALLY DIFFERENTIABLE *at x_0 in direction $0 \neq u \in X$ if there exist $\Delta_0 > 0$ and $\xi \in X^*$ such that for each $\varepsilon > 0$ there is $\Delta_1 = \Delta_1(\varepsilon) > 0$ for which φ is defined on $D_u(x_0, 2\Delta_0, \Delta_1)$ satisfying*

$$|\varphi(y) - \varphi(x) - \langle\xi, y - x\rangle| \leq \varepsilon\|y - x\| \;\; \textit{for all} \;\; x \in D_u(x_0, \Delta_0, \Delta_1),\; y \in D_u(x, \Delta_0, \Delta_1). \tag{4.2}$$

Observe that in the above definition we have $x \in D_u(x_0, \Delta_0, \Delta_1) \subset D_u(x_0, 2\Delta_0, \Delta_1)$, and hence the values $\varphi(x)$ in (4.2) are well defined. Similarly the values $\varphi(y)$ are well defined by $y \in D_u(x_0, 2\Delta_0, \Delta_1)$. Indeed, it follows that $y \in B(x_0, 2\Delta_0)$ due to the relationships

$$\|y - x_0\| \leq \|y - x\| + \|x - x_0\| \leq \Delta_0 + \Delta_0 = 2\Delta_0 \;\; \text{and} \;\; y \in x_0 + C_{u+\Delta_1 B},$$

where $y - x_0 = (y - x) + (x - x_0) \in C_{u+\Delta_1 B} + C_{u+\Delta_1 B} \subset C_{u+\Delta_1 B}$ by the convexity of $C_{u+\Delta_1 B}$.

The following theorem shows that directional subdifferentials of strictly directionally differentiable functions behave similarly to directional subdifferentials of linear continuous functions considered in Proposition 3.7.

Theorem 4.3 (directional subgradients of strictly directionally differentiable functions). *Let $\varphi\colon \operatorname{dom}\varphi \to \mathbb{R}$ be strictly differentiable at $x_0 \in \operatorname{dom}\varphi \subset X$ in direction $0 \neq u \in X$. Then we have the relationship*

$$\partial_u\varphi(x_0) = \big\{x^* \in X^* \big|\ \langle x^*, u\rangle \le \varphi'_u(x_0; u)\big\} \quad \text{with} \quad \varphi'_u(x_0; u) = \langle \xi, u\rangle, \tag{4.3}$$

where $\xi \in X^*$ is given in Definition 4.2 and satisfies $\xi \in \partial_u\varphi(x_0)$.

Proof. Taking $\Delta_0 > 0$ and $\xi \in X^*$, for every fixed $\varepsilon > 0$ find $\Delta_1 = \Delta_1(\varepsilon)$ such that (4.2) holds whenever $x \in D_u(x_0, \Delta_0, \Delta_1)$ and $y \in D_u(x, \Delta_0, \Delta_1)$. Let $x \in D_u(x_0, \Delta_0, \Delta_1)$, define $y := x + tv$ with $0 < t < \Delta_0/(\|u\| + \Delta_1)$, and select $v \in u + \Delta_1 B$. Hence $y \in D_u(x, \Delta_0, \Delta_1)$ and relationship (4.2) can be written as

$$\Big\|\frac{1}{t}\big(\varphi(x + tv) - \varphi(x)\big) - \langle \xi, u\rangle\Big\| \le \varepsilon\|u\|,$$

which implies that the directional derivative $\varphi'(x_0; u)$ exists and is computed by $\varphi'(x_0; u) = \langle \xi, u\rangle$.

To justify that $\xi \in \partial_u\varphi(x_0)$, pick a sequence $\varepsilon_k \to 0^+$ and choose $0 < \delta_k \le \Delta_1(\varepsilon_k)$ with $\delta_k \to 0^+$ as $k \to \infty$. Taking further $x_k \in D_u(x_0, \Delta_0, \delta_k)$ and $x \in D_u(x_k, \Delta_0, \delta_k)$, we get $x_k \overset{X,u}{\to} x_0$ as $k \to \infty$. Put $x_k^* := \xi \overset{w^*}{\to} \xi$ and observe that (3.3) follows from (4.2). Indeed, for $r \ge \varphi(x)$ we have the relationships

$$\begin{aligned}
&\langle x_k^*, x - x_k\rangle - \big(r - \varphi(x_k)\big) \le \langle \xi, x - x_k\rangle - \big(\varphi(x) - \varphi(x_k)\big)\\
&\le \big|\big(\varphi(x) - \varphi(x_k)\big) - \langle \xi, x - x_k\rangle\big| \le \varepsilon_k\|x - x_k\| \le \varepsilon_k\big(\|x - x_k\| + |r - \varphi(x_k)|\big),
\end{aligned}$$

which imply that $\xi \in \partial_u\varphi(x_0)$. Finally, the inclusion "$\subset$" in (4.3) follows from Theorem 3.6 and the opposite one "$\supset$" from Theorem 3.3. This completes the proof. $\square$

Next we establish a sum rule for the directional subdifferential (3.2) similar to the one in [11, Proposition 1.107] for the basic subdifferential (3.1) while replacing strictly differentiable functions by their directional counterparts.

Theorem 4.4 (sum rule for directional subgradients). *Let $\varphi\colon \Omega \to \mathbb{R}$ be strictly differentiable at $x_0 \in \Omega \subset X$ in direction $0 \neq u \in X$, and let $\psi\colon \Omega \to X$ be an arbitrary function finite at x_0. Then we have*

$$\partial_u(\varphi + \psi)(x_0) = \partial_u\varphi(x_0) + \partial_u\psi(x_0). \tag{4.4}$$

Proof. Take $\xi \in X^*$ from Definition 4.2 for φ at x_0. First we justify the inclusion "$\supset$" in (4.4). Fix arbitrary subgradients $x^* \in \partial_u\varphi(x_0)$ and $y^* \in \partial_u\psi(x_0)$. Applying the subgradient description (3.3) to the latter inclusion allows us to find sequences $\varepsilon_k \to 0^+$, $\delta_k \to 0^+$, $x_k \overset{\Omega,u}{\to} x_0$, and $y_k^* \overset{w^*}{\to} y^*$ as $k \to \infty$ such that whenever $\varepsilon > 0$ there are $\delta(k) > 0$ as $k \in \mathbb{N}$ with

$$\langle y_k^*, x - x_k\rangle - \big(r_\psi - \psi(x_k)\big) \le \Big(\varepsilon_k + \frac{\varepsilon}{4 + 2\|\xi\|}\Big)\big(\|x - x_k\| + |r_\psi - \psi(x_k)|\big)$$

for all $r_\psi \ge \psi(x)$ and $x \in D_u(x_k, \delta(k), \delta_k)$. The obtained inequality yields for sufficiently large $k \in \mathbb{N}$ the following chain of relationships:

$$
\begin{aligned}
&\langle y_k^*, x - x_k\rangle - \big(r_\psi - \psi(x_k)\big)\\
&\le \Big(\varepsilon_k + \frac{\varepsilon}{4+2\|\xi\|}\Big)\big(\|x - x_k\| + |(r_\psi + \varphi(x)) - (\varphi(x_k) + \psi(x_k))| + |\varphi(x) - \varphi(x_k)|\big)\\
&\le \Big(\varepsilon_k + \frac{\varepsilon}{4+2\|\xi\|}\Big)\big(\|x - x_k\| + |(r_\psi + \varphi(x)) - (\varphi(x_k) + \psi(x_k))| + \|x - x_k\| + |\langle \xi, x - x_k\rangle|\big)\\
&\le \Big(\varepsilon_k(2 + \|\xi\|) + \frac{\varepsilon}{2}\Big)\big(\|x - x_k\| + |(r_\psi + \varphi(x)) - (\varphi(x_k) + \psi(x_k))|\big).
\end{aligned}
$$

Pick $r \ge \varphi(x) + \psi(x)$ and denote $r_\psi := r - \varphi(x) \ge \psi(x)$, which gives $r - r_\psi \ge \varphi(x)$. As follows from the proof of Theorem 4.3, for $x_k^* := x^* \xrightarrow{w^*} x^*$ and sufficiently large $k \in \mathbb{N}$ we have

$$
\begin{aligned}
&\langle x_k^*, x - x_k\rangle - (r - r_\psi - \varphi(x_k))\\
&\le \frac{\varepsilon}{2}\|x - x_k\| \le \Big(\frac{\varepsilon}{2}(\|x - x_k\| + |r - (\varphi(x_k) + \psi(x_k))|\Big).
\end{aligned} \tag{4.5}
$$

Adding (4.5) to the estimate above ensures that

$$
\begin{aligned}
&\langle x_k^* + y_k^*, x - x_k\rangle - \big(r - (\varphi(x_k) + \psi(x_k))\big)\\
&\le \big(\varepsilon_k(2 + \|\xi\|) + \varepsilon\big)\big(\|x - x_k\| + |r - (\varphi(x_k) + \psi(x_k))|\big).
\end{aligned} \tag{4.6}
$$

Since $x_k^* + y_k^* \xrightarrow{w^*} x^* + y^*$, we get from (4.6) and the subgradient description (3.3), with $\varepsilon_k(2 + \|\xi\|) \to 0^+$ replacing ε_k, that $x^* + y^* \in \partial_u(\varphi + \psi)(x_0)$. This justifies the inclusion "$\supset$" in (4.4).

To prove the opposite inclusion "$\subset$" in (4.4), observe that the function $-\varphi$ is strictly differentiable at x_0 in direction u, with $-\xi \in X^*$ instead of ξ in Definition 4.2. It follows from Theorem 4.3 that $-\xi \in \partial_u(-\varphi)(x_0)$. Picking an arbitrary subgradient $z^* \in \partial_u(\varphi + \psi)(x_0)$ and employing the above arguments, we get $z^* - \xi \in \partial_u((\varphi + \psi) - \varphi)(x_0) = \partial_u\psi(x_0)$ and thus $z^* = \xi + (z^* - \xi) \in \partial_u\varphi(x_0) + \partial_u\psi(x_0)$, which completes the proof of the theorem. □

Let us consider an important class of nondifferentiable functions strictly directionally differentiable at their characteristic points in the sense of Definition 3.5 and compute their directional subgradients. Note that the proof given below is substantially more involved in comparison with standard directional derivatives due to more subtle definition of our construction motivated by advanced applications to optimization.

Proposition 4.5 (strict directional differentiability of norm functions). *In a Banach space X fix a point $x_0 \in X$ and consider the norm function $\varphi(x) := \|x - x_0\|$. Then it is strictly differentiable at x_0 in any direction $0 \neq u \in X$ and its directional subgradients $x^* \in \partial_u\varphi(x_0)$ at x_0 are computed by $\langle x^*, u\rangle \le \langle \xi, u\rangle$ with any $\xi \in X^*$ such that $\langle \xi, u\rangle = \|u\|$ with $\|\xi\| = 1$.*

Proof. Observe first that the existence of elements $\xi \in X^*$ mentioned in the proposition follows from the classical Hahn-Banach theorem. We are going to show that all such elements ξ are directional subgradients of φ at the underlying point x_0. To proceed, choose $0 < \delta < \|u\|$, $x \in x_0 + C_{u+\delta}$, and $y \in x + C_{u+\delta}$. It follows from (3.6) that $x - x_0 = \lambda_x(u + \delta b_x)$ and $y - x = \lambda_y(u + \delta b_y)$ for some elements $b_x, b_y \in B$, $\lambda_x := \|x - x_0\|/(\|u\| + \theta_x)$ with $-\delta \le \theta_x \le \delta$, and $\lambda_y := \|y - x\|/(\|u\| + \theta_y)$ with $-\delta \le \theta_y \le \delta$.

Put $\overline{y} := x + \lambda_y(u + \delta b_x)$ and hence get $\overline{y} = x_0 + (\lambda_x + \lambda_y)(u + \delta b_x)$ with $\|\overline{y} - x_0\| = (\lambda_x + \lambda_y)\|u + \delta b_x\|$. It also follows that $y = x_0 + \lambda_x(u + \delta b_x) + \lambda_y(u + \delta b_y)$ with $y - \overline{y} = (y - x_0) - (\overline{y} - x_0) = \lambda_y \delta(b_y - b_x)$. The triangle inequality gives the estimates

$$\begin{aligned} \|y - x_0\| = \|(\overline{y} - x_0) + (y - \overline{y}))\| \leq (\lambda_x + \lambda_y)\|u + \delta b_x\| + 2\lambda_y\delta, \\ \|y - x_0\| = \|(\overline{y} - x_0) + (y - \overline{y}))\| \geq (\lambda_x + \lambda_y)\|u + \delta b_x\| - 2\lambda_y\delta. \end{aligned}$$

Taking into account that $\lambda_x\|u + \delta b_x\| = \|x - x_0\|$, the latter inequalities can be rewritten as

$$\lambda_y\|u + \delta b_x\| - 2\lambda_y\delta \leq \|y - x_0\| - \|x - x_0\| \leq \lambda_y\|u + \delta b_x\| + 2\lambda_y\delta. \tag{4.7}$$

Let $\xi \in X^*$ and observe the representations

$$-\langle \xi, y - x\rangle = -\langle \xi, \lambda_y(u + \delta b_y)\rangle = -\lambda_y\langle \xi, u\rangle - \lambda_y\delta\langle \xi, b_y\rangle,$$

which ensure in turn the estimates

$$-\lambda_y\langle \xi, u\rangle - \lambda_y\delta\|\xi\| \leq -\langle \xi, y - x\rangle \leq -\lambda_y\langle \xi, u\rangle + \lambda_y\delta\|\xi\|. \tag{4.8}$$

Summing up the inequalities in (4.7) and (4.8) gives us

$$\begin{aligned} \lambda_y(\|u + \delta b_x\| - \langle \xi, u\rangle) - \lambda_y(2 + \|\xi\|)\delta \leq \|y - x_0\| - \|x - x_0\| - \langle \xi, y - x\rangle \\ \leq \lambda_y(\|u + \delta b_x\| - \langle \xi, u\rangle) + \lambda_y(2 + \|\xi\|)\delta. \end{aligned}$$

Letting now $\langle \xi, u\rangle = \|u\|$, we get

$$-\delta \leq \|u + \delta b_x\| - \langle \xi, u\rangle = \|u + \delta b_x\| - \|u\| \leq \delta$$

and arrive therefore at the estimates

$$-\lambda_y(3 + \|\xi\|)\delta \leq \|y - x_0\| - \|x - x_0\| - \langle \xi, y - x\rangle \leq -\lambda_y(3 + \|\xi\|)\delta.$$

If $x \in D_u(x_0, \Delta_0, \Delta_1)$ and $y \in D_u(x_0, \Delta_0, \Delta_1)$ with $0 < \Delta_1 < \|u\|$ and $\lambda_y \leq \|y - x\|/(\|u\| - \Delta_1) \leq \Delta_0/(\|u\| - \Delta_1)$, we obtain from the above that

$$|\varphi(y) - \varphi(x) - \langle \xi, y - x\rangle| \leq \frac{(3 + \|\xi\|)\Delta_0\Delta_1}{\|u\| - \Delta_1} < \varepsilon \ \text{ for } \ \Delta_1 < \min\Big(\frac{\varepsilon\|u\|}{\Delta_0(3 + \|\xi\|) + \varepsilon}, \|u\|\Big),$$

which justifies the strict directional differentiability of the norm function and $\xi \in \partial_u\varphi(x_0)$. If $\langle x^*, u\rangle \leq \langle \xi, u\rangle$ then $x^* \in \partial_u\varphi(x_0)$ on the basis of Theorem 3.3 and thus completes the proof of the proposition. □

It is well known that a function φ strictly differentiable at x_0 is locally Lipschitzian around this point. It is natural to introduce a directional extension of locally Lipschitzian functions, which encompasses strictly directionally differentiable ones. Note the class of functions locally Lipschitzian in directions introduced in the next definition is essentially different from the class of directionally Lipschitzian functions in the sense of Rockafellar [14].

Definition 4.6 (real-valued functions Lipschitzian in directions). *A function $\varphi\colon \Omega \to \mathbb{R}$ is* LOCALLY LIPSCHITZIAN *around $x_0 \in \Omega \subset X$* IN DIRECTION *$0 \neq u \in X$ with constant $\ell \geq 0$ if there are numbers $\Delta_0 > 0$ and $\Delta_1 > 0$ such that*

$$|\varphi(y) - \varphi(x)| \leq \ell\|y - x\| \ \textit{ for all } \ x \in \Omega \cap D_u(x_0, \Delta_0, \Delta_1) \ \textit{ and } \ y \in \Omega \cap D_u(x, \Delta_0, \Delta_1).$$

It is easy to observe that strictly directionally differentiable functions satisfy the directional Lipschitzian property of Definition 4.6.

Proposition 4.7 (strictly directionally differentiable functions are Lipschitzian in directions). *Let $\varphi\colon \Omega \to \mathbb{R}$ be strictly differentiable at x_0 in direction $0 \neq u \in X$. Then it is locally Lipschitzian around x_0 in this direction.*

Proof. According to the strict directional differentiability (4.2), for all $x \in D_u(x_0, \Delta_0, \Delta_1)$ and $y \in D_u(x, \Delta_0, \Delta_1)$ we have the estimates

$$\begin{aligned}|\varphi(y) - \varphi(x)| &\leq |\varphi(y) - \varphi(x) - \langle \xi, y - x\rangle| + |\langle \xi, y - x\rangle| \\ &\leq \varepsilon\|y - x\| + \|\xi\| \cdot \|y - x\| \leq (\varepsilon + \|\xi\|)\|y - x\|.\end{aligned}$$

By Definition 4.6 the latter means that the function φ is locally Lipschitzian around x_0 with constant $\ell = \varepsilon + \|\xi\|$ in direction u under consideration. □

To conclude this section, observe that, in contrast to the classical definitions of strict differentiability and local Lipschitz continuity involving a neighborhood $B(x_0, \Delta_0)$ of the reference point, our directional constructions are well-defined and provide significant information at boundary points, which is essential for applications to optimization.

5 Unconstrained Optimization

In this section we start the study of optimization problems by applying the tools of directional generalized differentiation developed above as well as their appropriate modifications. Our first attention is paid to problems of unconstrained optimization by which we understand minimizing a function $f\colon \Omega \to \mathbb{R}$ defined on a given set Ω with no additional constraints. Note that in this problem we treat the set $\Omega \subset X$ not as a constraint but rather as a domain region, which may not coincide with the whole linear space X. Some of the results below deal with the case of $\Omega = X$.

We begin with the following directional subdifferential counterpart of the Fermat stationary principle obtained in terms of our basic directional subdifferential (3.2).

Theorem 5.1 (generalized Fermat principle via basic directional subgradients). *Let $x_0 \in \Omega \subset X$ be a local minimizer (with respect to Ω) of a function $f\colon \Omega \to \mathbb{R}$ on a given subset Ω of a Banach space X. Then for any direction $0 \neq u \in X$ we have the inclusion*

$$0 \in \partial_u f(x_0). \tag{5.1}$$

Proof. Using the subgradient description (3.3), put therein $x_k := x_0 \xrightarrow{\Omega,u} x_0$, $x_k^* := 0 \xrightarrow{w^*} x^* = 0$ and pick $\varepsilon_k \to 0^+$ and $\delta_k \to 0^+$ arbitrarily. Then we have (5.1) directly from (3.3) since the left-hand side of the latter inequality is nonpositive. □

Next we explore the possibility to obtain some other versions of the directional Fermat principle (5.1) by employing two modifications (in fact simplifications) of the basic directional subdifferential (3.2), which have similar bur somewhat different properties and applications. The first modification is defined as follows.

(M1) Modify definition (2.7) of the limiting normal cone with respect to a set by taking in (2.7) the convergence $x \overset{\Omega\setminus\{x_0\},u}{\to} x_0$ instead of $x \overset{\Omega,u}{\to} x_0$ if $u \in T(x_0,\Omega)$ and putting X^* otherwise. Note that the convergence $x \overset{\Omega\setminus\{x_0\},u}{\to} x_0$ means $x \to x_0$ with $x \neq x_0$ and such that $\frac{x-x_0}{\|x-x_0\|} \to \frac{u}{\|u\|}$. Construct further the directional subdifferential $\partial_u\varphi(x_0)$ as in Definition 3.1 with this new meaning for the normal cone $N_{u\times\mathbb{R}}((x_0,\varphi(x_0));\operatorname{epi}\varphi)$.

The following example shows that Theorem 5.1 formulated with this modification is not true in general; see however Theorem 5.3.

Example 5.2 (violation of the directional Fermat principle via modified subgradients from (M1)). Define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x = 0, \\ 1 & \text{if } x \text{ is irrational}, \\ \dfrac{1}{q} & \text{if } x \text{ is rational with } |x| = \dfrac{p}{q}, \ p \text{ and } q \text{ relatively prime.} \end{cases}$$

Obviously $x_0 = 0$ is a global minimizer of f on $\mathbb{R}$. At the same time there is no $x_k \overset{\mathbb{R}\setminus\{0\},1}{\to} 0$ for which (3.3) holds, and thus inclusion (5.1) is violated for the subdifferential modification in (M1). To justify this, take any sequence $x_k \to 0$ with $x_k \neq 0$, divide it by $|x|$, and observe that

$$f'_-(x_k;1) := \liminf_{x\to x_k^+} \frac{f(x)-f(x_k)}{x-x_k} = -\infty.$$

Nevertheless the next theorem shows that the directional Fermat principle (5.1) holds in terms of the modified directional subdifferential from (M1) under some additional assumptions.

Theorem 5.3 (directional Fermat principle via modified subgradients from (M1)). *Let $f\colon \Omega \to \mathbb{R}$ be lower semicontinuous (l.s.c.) on a subset $\Omega \subset X$ in finite dimensions, and let $x_0 \in \Omega$ be a local minimizer of f on Ω. Assume that Ω is locally closed around x_0 and that f is continuous at x_0. Then* (5.1) *holds in terms of the modified directional subdifferential of f at x_0 from* (M1).

Proof. If $u \notin T(x_0,\Omega)$, the conclusion of the theorem is obvious since $\partial_u f(x_0) = X^*$. Considering now the case of $u \in T(x_0,\Omega)$, we justify the following

Claim. *For all $\Delta_0,\Delta_1 > 0$ there exists $x \in (\Omega\setminus\{x_0\})\cap\overline{D}_u(x_0,\Delta_0,\Delta_1)$ with $\|x-x_0\| < \Delta_0$ such that whenever $\varepsilon > 0$ there is $\delta > 0$ depending also on ε with the properties:*

$$\overline{D}_u(x,\delta(x),\Delta_1) \subset \overline{D}_u(x_0,\Delta_0,\Delta_1), \tag{5.2}$$

$$f(y) - f(x) \geq -\varepsilon\|y-x\| \quad \textit{for all} \quad y \in \Omega\cap\overline{D}_u(x,\delta(x),\Delta_1). \tag{5.3}$$

To prove this claim, we argue by contradiction. Select numbers $\Delta_0,\Delta_1 > 0$ arbitrarily while make Δ_0 sufficiently small so that the set $\Omega\cap\overline{B}(x_0,\Delta_0)$ is closed and $f(x_0) \leq f(x)$ for all $x \in \Omega\cap\overline{B}(x_0,\Delta_0)$. Furthermore, Δ_1 is sufficiently small to sure that the relationships $x \in x_0 + C_{u+\Delta_1\overline{B}}$,

$y \in x + C_{u+\Delta_1\overline{B}}$, and $y \neq x$ imply that $\|y - x_0\| > \|x - x_0\|$. The latter can be done by the following reasons. Taking

$$
\begin{aligned}
x - x_0 = \lambda_x(u + \Delta_1 b_x), \quad b_x \in \overline{B}, \quad \frac{\|x-x_0\|}{\|u\|+\Delta_1} \le \lambda_x \le \frac{\|x-x_0\|}{\|u\|-\Delta_1},\\
y - x = \lambda_y(u + \Delta_1 b_y), \quad b_y \in \overline{B}, \quad \frac{\|y-x\|}{\|u\|+\Delta_1} \le \lambda_y \le \frac{\|y-x\|}{\|u\|-\Delta_1},
\end{aligned}
$$

we arrive at the relationships

$$
\begin{aligned}
&\|y - x_0\| = \|(y-x) + (x-x_0)\| = \|(\lambda_x+\lambda_y)u + \Delta_1(\lambda_x b_x + \lambda_y b_y))\|\\
&\ge (\lambda_x+\lambda_y)\|u\| - \Delta_1(\lambda_x+\lambda_y) \ge \left(\frac{\|x-x_0\|}{\|u\|+\Delta_1} + \frac{\|y-x\|}{\|u\|+\Delta_1}\right)\|u\|\\
&-\Delta_1\left(\frac{\|x-x_0\|}{\|u\|+\Delta_1} + \frac{\|y-x\|}{\|u\|+\Delta_1}\right) \overset{\Delta_1\to 0^+}{\to} \|x-x_0\| + \|y-x\| > \|x-x_0\|.
\end{aligned}
$$

To proceed, for $x \in (\Omega\setminus\{x_0\}) \cap \overline{D}_u(x_0,\Delta_0,\Delta_1)$ with $\|x - x_0\| < \Delta_0$ denote by $\mathcal{D}(x)$ the set points

$$
y \in \Omega \cap (x + C_{u+\Delta_1\overline{B}}) \cap \overline{D}_u(x_0,\Delta_0,\Delta_1) \tag{5.4}
$$

satisfying the condition

$$
f(y) - f(x) \le -\varepsilon\|y - x\|. \tag{5.5}
$$

By the assumption made we have $\mathcal{D}(x) \neq \emptyset$. Let us show that the set $\mathcal{D}(x)$ is closed. Indeed, consider a sequence $y_k \to y$ as $k \to \infty$ with $y_k \in \mathcal{D}(x)$ for all $k \in \mathbb{N}$. Then inclusion (5.4) holds for the limiting point y due to the closedness of the set on the right-hand side of (5.4). Condition (5.5) also holds for such y by the lower semicontinuity of f. Since X is finite-dimensional, the set $\mathcal{D}(x)$ is compact. Picking further $\overline{y} \in \mathcal{D}(x)$ so that $\|\overline{y} - x_0\| = \sup\{\|y - x_0\| \mid y \in \mathcal{D}(x)\}$ and prove that $\|\overline{y} - x_0\| = \Delta_0$. If this is not the case, find $y \in \mathcal{D}(\overline{y})$ and observe that it yields the inclusion $y \in \mathcal{D}(x)$, which follows from

$$
y \in \overline{y} + C_{u+\Delta_1\overline{B}} \subset x + C_{u+\Delta_1\overline{B}} + C_{u+\Delta_1\overline{B}} = x + C_{u+\Delta_1\overline{B}} \tag{5.6}
$$

by taking into account that $C_{u+\Delta_1\overline{B}}$ is a convex cone. We also have

$$
f(y) - f(x) = (f(y) - f(\overline{y})) + (f(\overline{y}) - f(x)) \le -\varepsilon(\|y - \overline{y}\| + \|\overline{y} - x\|) \le -\varepsilon\|y - x\| \tag{5.7}
$$

by the triangle inequality. It follows from the choice of Δ_1 that $\|y - x_0\| > \|\overline{y} - x_0\|$, which contradicts the choice of $\overline{y}$.

To complete the proof of the claim, consider now a sequence of $x_k \in (\Omega\setminus\{x_0\}) \cap \overline{D}_u(x_0,\Delta_0,\Delta_1)$ with $x_k \overset{\Omega,u}{\to} x_0$ as $k \to \infty$; such a choice is possible due to $u \in T(x_0,\Omega)$. Then pick $y_k \in \mathcal{D}(x_k)$ with $\|y_k - x_0\| = \Delta_0$ and assume without loss of generality that $y_k \to y_0$ as $k \to \infty$. Passing finally to the limit in the inequality $f(y_k) - f(x_k) \le -\varepsilon\|y_k - x_k\|$ with taking into account that f is assumed to be continuous at x_0 and l.s.c. at y_0 gives us

$$
f(y_0) - f(x_0) \le \liminf_{k\to\infty}\big(f(y_k) - f(x_k)\big) \le -\varepsilon\|y_0 - x_0\| = -\varepsilon\Delta_0 < 0, \tag{5.8}
$$

which contradicts the local minimality of x_0 and thus justifies the claim.

Now to finish the proof of the theorem, let $\varepsilon > 0$ and let $\Delta_0 > 0$ be such that the set $\Omega \cap \overline{B}(x_0,\Delta_0)$ is closed and that $f(x_0) \le f(x)$ for $x \in \Omega \cap \overline{B}(x_0,\Delta_0)$. Take a sequence of $\gamma_k \le \Delta_0$

with $\gamma_k \to 0^+$ as $k \to \infty$. Letting $\delta_k \to 0^+$, choose $x_k \in (\Omega \setminus \{x_0\}) \cap \overline{D}_u(x_0, \gamma_k, \delta_k)$ satisfying $\overline{D}_u(x_k, \delta(x_k), \delta_k) \subset \overline{D}_u(x_0, \gamma_k, \delta_k)$ along some $\delta(x_k)$ and such that

$$f(x) - f(x_k) \geq -\varepsilon \|x - x_k\| \quad \text{whenever} \quad x \in \overline{D}_u(x_k, \delta(x_k), \delta_k).$$

Since $\gamma_k \to 0^+$ and $\delta_k \to 0^+$, we have $x_k \overset{\Omega,u}{\to} x_0$ as $k \to \infty$. Choose $\varepsilon_k \to 0^+$ arbitrary and put $x_k^* := 0$ and $x^* := 0$; hence $x_k^* \to x^*$. Finally, for all $x \in D_u(x_k, \delta(x_k), \delta_k)$ and $r \geq f(x)$ we have

$$\begin{aligned} &\langle x_k^*, x - x_k \rangle - \big(r - f(x_k)\big) \leq -\big(f(x) - f(x_k)\big) \\ &\leq \varepsilon \|x - x_k\| \leq (\varepsilon_k + \varepsilon)\big(\|x - x_k\| + |r - f(x_k)|\big), \end{aligned}$$

which shows that $0 \in \partial_u f(x_0)$ for the directional subdifferential modification from (M1) and thus completes the proof of the theorem. $\square$

Next we consider yet another modification of the our basic directional subdifferential (3.2) and explore its possible applications.

(M2) Modify definition (2.7) of the limiting normal cone with respect to a set taking on the right-hand side of this construction the collection of ε-normals, as in the basic normal cone case (2.3), instead of the ε-normals with respect to a set. Furthermore, similarly to the modification in (M1), include only $x \neq x_0$ in the limiting procedure and thus arrive at the construction

$$N_Q(x_0; \Omega) := \operatorname*{Lim\,sup}_{x \overset{\Omega \setminus \{x_0\}, Q}{\to} x_0, \, \varepsilon \to 0+} \widehat{N}_\varepsilon(x; \Omega). \tag{5.9}$$

Then define $\partial_u \varphi(x_0)$ as in (3.2) with this new meaning (5.9) for the cone $N_{u \times \mathbb{R}}((x_0, \varphi(x_0)); \operatorname{epi} \varphi)$.

There are some reasons to define directional subdifferentials according to (M2), which is actually done in our previous paper [7]. If the normal cone with respect to a set is defined via the ε-normals as (5.9), it is easier to see relations of it with the basic normal cone (2.3). Such relationships in this vein are given in [7]. Excluding x_0 when x tends directionally to x_0, we eliminate the influence of ε-normals at the reference point x_0, which show rather "nondirectional behavior". Still the directional subdifferential modified in this way may have some advantages with respect to the basic nondirectional subdifferential (3.1). As shown in [7], necessary optimality conditions in terms of (M2) can be effective while similar optimality conditions in terms of the basic subdifferential (3.1) fail. We have the following result proved in [7] in a more general setting of constrained optimization.

Theorem 5.4 (scalarized necessary optimality conditions). *Let $f : X \to \mathbb{R}$ be locally Lipschitzian around its local minimizers x_0, and let $\dim X < \infty$. Then we have*

$$\langle \partial_u f(x_0), u \rangle \cap \mathbb{R}_+ \neq \emptyset \quad \text{for all} \quad 0 \neq u \in X, \tag{5.10}$$

where the directional subdifferential $\partial_u f(x_0)$ is defined as in (M2).

Observe that the proof of Theorem 5.4 given in [7] are based on integration with respect to the Lebesgue measure and the Rademacher theorem for locally Lipschitzian functions in finite dimensions. The following example shows the local Lipschitz continuity of the cost function is essential for the validity of the scalarized necessary optimality conditions (5.10).

Example 5.5 (violation of the scalarized optimality condition for non-Lipschitzian functions). Consider $X = \mathbb{R}$ and the function $f(x) := \sqrt{|x|}$, which attains its global minimum at $x_0 = 0$. It is easy to check that $\partial_u f(x_0) = \emptyset$ when $u = \pm 1$ for the subdifferential construction of Theorem 5.4, and thus and necessary optimality condition condition (5.10) does not hold. Observe, that for this function the directional Fermat principle (5.1) of Theorem 5.1 is satisfied.

Another natural question arises about the validity of the stationary condition $0 \in \partial_u f(x_0)$ via the modified directional subdifferential from (M2) in the framework of Theorem 5.4. The next example demonstrates that such a version of the directional Fermat principle from Theorem 5.1 does not hold even for Lipschitzian function on the real line.

Example 5.6 (violation of the directional Fermat principle via the modified subdifferential from (M2)). Consider a real function $f(x) := |x|$ on $\mathbb{R}$, which attains its minimum at $x_0 = 0$. It is easy to calculate the directional subdifferential from (M2) as $\partial_u f(x_0) = 1$ for $u = 1$ and $\partial_u f(x_0) = -1$ for $u = -1$. Condition (5.10) is satisfied while that of $0 \in \partial_u f(x_0)$ fails for the directional subdifferential modified in (M2). Observe that this does not contradict the directional Fermat principle of Theorem 5.1 in terms of our basic directional subdifferential (3.2) for which we have $\partial_u f(x_0) = (-\infty, 1]$ when $u = 1$ and $\partial_u f(x_0) = [-1, \infty)$ when $u = -1$ (cf. Proposition 4.7 and Theorem 4.3), i.e., in both cases the inclusion $0 \in \partial_u f(x_0)$ holds.

Next we establish a sufficient optimality condition in an appropriate modification of form (5.10) via the directional subdifferential modified in (M2). Recall that $x_0 \in \Omega$ is a *first-order isolated minimizer* for $f \colon \Omega \to \mathbb{R}$ with $\Omega \subset X$ if there are constants $r > 0$ and $\alpha > 0$ such that

$$f(x) - f(x_0) \geq \alpha \|x - x_0\| \quad \text{whenever} \quad x \in \Omega \cap B(x_0, r). \tag{5.11}$$

In the proof of the sufficient optimality condition we use the following lemma taken from [7].

Lemma 5.7 (mean value estimates). *Let the function $\varphi \colon D_u(x_0, \alpha, \beta) \to \mathbb{R}$ be Lipschitz continuous on the set $D_u(x_0, \alpha, \beta) \subset X := \mathbb{R}^n$ with some $x_0 \in X$, $u \in X \setminus \{0\}$, and α, $\beta > 0$. Suppose that there exists $c \in \mathbb{R}$ such that*

$$\langle \varphi'(x), u \rangle \leq c \quad (\textit{resp.} \ \ \langle \varphi'(x), u \rangle \geq c)$$

holds almost everywhere with respect to the n-dimensional Lebesgue measure λ_n. Then for all $t \in [0, \alpha]$ we have the estimate

$$\varphi(x_0 + tu) - \varphi(x_0) \leq ct \quad (\textit{resp.} \ \ \varphi(x_0 + tu) - \varphi(x_0) \geq ct).$$

Here is the aforementioned sufficient condition for first-order isolated minimizers.

Theorem 5.8 (sufficient optimality condition for isolated minimizers). *Let $f \colon X \to \mathbb{R}$ be locally Lipschitzian around x_0, and let $\dim X < \infty$. Suppose that*

$$\langle \partial_u f(x_0), u \rangle \subset \operatorname{int} \mathbb{R}_+ \quad \textit{for all} \ \ 0 \neq u \in X \tag{5.12}$$

via the directional subdifferential $\partial_u f(x_0)$ modified as in (M2). *Then x_0 is a first-order isolated minimizer for the function f on X.*

Proof. Denote by $\ell \geq 0$ a Lipschitz constant of f on some ball $B(x_0, r)$ and assume on the contrary that (5.11) does not hold. The latter means that for any $\varepsilon_k \to 0^+$ there is a sequence $x_k \to x_0$ as $k \to \infty$ with $x_k \neq x_0$ such that

$$f(x_k) - f(x_0) \leq \varepsilon_k \|x_k - x_0\| \quad \text{for all} \quad k \in \mathbb{N}. \tag{5.13}$$

Passing to a subsequence by the compactness of the unit sphere in finite dimensions, we suppose that $x_k \overset{X,u}{\to} x_0$ as $k \to \infty$ for some $u \in X$ with $\|u\| = 1$. Applying inclusion (5.12) to this u together with the construction of $\partial_u f(x_0)$ and the classical Rademacher theorem, we find constants $c > 0$, $\alpha > 0$, and $0 < \beta < r$ such that on the set $D_u(x_0, \alpha, \beta)$ the following holds: if the derivative $f'(x)$ exists at some $x \in D_u(x_0, \alpha, \beta)$, then $\langle f'(x), u\rangle \geq c$. By Lemma 5.7 the latter implies that

$$f(x_0 + tu) - f(x_0) \geq ct \quad \text{for all} \quad 0 \leq t \leq \alpha. \tag{5.14}$$

Putting $t := \|x_k - x_0\|$, we get from (5.14) and the Lipschitz property of f that

$$\begin{aligned}
\ell\|x_k - x_0\| \cdot \left\| \frac{x_k - x_0}{\|x_k - x_0\|} - u \right\| &\geq f(x_k + \|x_k - x_0\|u) - f(x_k) \\
&= \big(f(x_k + \|x_k - x_0\|u) - f(x_0)\big) - \big(f(x_k) - f(x_0)\big) \\
&\geq c\|x_k - x_0\| - \varepsilon_k\|x_k - x_0\| \quad \text{for all} \quad k \in \mathbb{N}.
\end{aligned}$$

Dividing this by $\|x_k - x_0\|$ gives us the estimate

$$\ell\left\| \frac{x_k - x_0}{\|x_k - x_0\|} - u \right\| \geq c - \varepsilon_k, \quad k \in \mathbb{N},$$

which implies in turn by passing to the limit as $k \to \infty$ that $0 \geq c$. The obtained contradiction completes the proof of the theorem. $\square$

Observe that, although the sufficient condition (5.12) is in the vein of the necessary condition (5.10), the following relationship

$$\langle \partial_u f(x_0), u\rangle \cap \operatorname{int} \mathbb{R}_+ \neq \emptyset \tag{5.15}$$

seems to be even closer to (5.10) than (5.12). The question is whether Theorem 5.8 remains true with the replacement (5.10) by (5.15) as a *necessary* condition for optimality. Another interesting question is about the possibility to replace (5.12) by the inclusion

$$\langle \partial_u f(x_0), u\rangle \subset \mathbb{R}_+ \tag{5.16}$$

as a *sufficient* optimality condition in the framework of Theorem 5.8. The answer is negative for both questions, which is demonstrated by the following counterexamples.

Example 5.9 (counterexamples for modified necessary and sufficient optimality conditions). Consider the sets on the real line

$$A_+ := \{0\} \cup \bigcup_{k=-\infty}^{\infty} [4^k, 3 \cdot 4^k] \quad \text{and} \quad A_- := \bigcup_{k=-\infty}^{\infty} (3 \cdot 4^k, 4^{k+1}).$$

Employing the characteristic functions χ_{A_-} and χ_{A_+} of these sets, define $\varphi_-, \varphi_+ : \mathbb{R} \to \mathbb{R}$ by

$$\varphi_-(x) := \int_0^{|x|} \chi_{A_-}(t)\, dt \quad \text{and} \quad \varphi_+(x) := \int_0^{|x|} \chi_{A_+}(t)\, dt.$$

(i) To show first that condition (5.15) is not necessary for optimality, consider the real function $f(x) = \varphi_+(x) - \varphi_-(x)$, which is obviously Lipschitz continuous on $\mathbb{R}$ with constant $\ell = 1$ with the property $f(x) \geq \frac{2}{3}|x|$. The point $x_0 = 0$ is a global minimizer of f; more precisely, even a first-order isolated minimizer of this function. For both $u = 1$ and $u = -1$ we get $\partial_u f(x_0) = [-1, 1]$. Therefore f satisfies the necessary condition (5.10) of Theorem 5.4 while does not satisfy (5.15). Observe that this examples shows also that condition (5.12) is not necessary for x_0 to be a first-order isolated minimizer for problems of unconstrained optimization.

(ii) Consider the function $g(x) := -f(x)$, where f is defined in part (i) of this example. It is easy to see that $\partial_u g(x_0) = [-1, 1]$ for x_0 and $u = \pm 1$. Thus condition (5.15) holds, but x_0 is not a minimizer of g, i.e,, definitely not a first-order isolated minimizer of this function as in Theorem 5.8. It is in a fact a maximizer of g.

6 Constrained Optimization

This section we study the following problem of constrained optimization with finitely many inequality constraints: given a cost function $f \colon \Omega \to \mathbb{R}$ defined on a set $\Omega \subset X$ and given constraint function $g_j \colon \Omega \to \mathbb{R}$ as $j = 1, \ldots, p$,

$$\text{minimize } \ f(x) \ \text{ subject to } \ g_j(x) \leq 0, \quad j = 1, \ldots, p. \tag{6.1}$$

The main attention is paid in this section to deriving necessary optimality conditions for local minimizers (with respect to the given domain set Ω) of problem (6.1) in terms of our basic directional subdifferential (3.2) of the cost and constraint functions. We discuss some motivating and illustrating examples and show that the new directional results may be more efficient that those expressed via the Mordukhovich limiting subdifferential (3.1) and other nondirectional constructions. Here we confine our consideration to the case of finite-dimensional spaces X, which seems to be more appropriate at the current stage of research to take advantages of directional subdifferential constructions incorporating the contingent cone (2.5).

First we recall the *prototyped* necessary optimality conditions obtained in [12, Theorem 5.19] for a local minimizer x_0 of problem (6.1) with functions f and g_j locally Lipschitzian around x_0: there are multipliers λ and $(\mu_1, \ldots, \mu_p)$, not all zero, such that

$$\lambda \geq 0, \ \ \mu_j \geq 0 \ \text{ with } \ \mu_j \, g_j(x_0) = 0 \ \text{ for } \ j = 1, \ldots, p, \ \text{ and} \tag{6.2}$$

$$0 \in \partial\Big(\lambda f + \sum_{j=1}^{p} \mu_j g_j\Big)(x_0). \tag{6.3}$$

The directional analogue of (6.3) in terms of the basic directional subdifferential (3.2) is

$$0 \in \partial_u\Big(\lambda f + \sum_{j=1}^{p} \mu_j g_j\Big)(x_0) \ \text{ for any } \ 0 \neq u \in X. \tag{6.4}$$

Note that we would like to be able selecting multipliers λ and μ_j *uniformly in all the directions* $0 \neq u \in X$ in the desired directional condition (6.4). The major question we discuss in this section is whether condition (6.4) combined with (6.2) is necessary for the local optimality of x_0

in problem (6.1) under some reasonable assumptions. To proceed, first we examine in detail the following motivating example, which sheds light on deriving a general result in this direction.

Example 6.1 (motivating example for directional subdifferential optimality conditions). Consider the constrained optimization problem (6.1) with $X = \mathbb{R}$, $p = 1$, and the functions $f\colon X \to \mathbb{R}$ and $g\colon X \to \mathbb{R}$ given as follows. Construct the sets

$$A_1 := \{0\} \cup \bigcup_{k=-\infty}^{\infty}[5^k, 2\cdot 5^k), \quad A_2 := \bigcup_{k=-\infty}^{\infty}[2\cdot 5^k, 3\cdot 5^k),$$
$$A_3 := \bigcup_{k=-\infty}^{\infty}[3\cdot 5^k, 4\cdot 5^k), \quad A_4 := \bigcup_{k=-\infty}^{\infty}[4\cdot 5^k, 5^{k+1})$$

and then define, via the characteristic functions of A_i, the following real functions

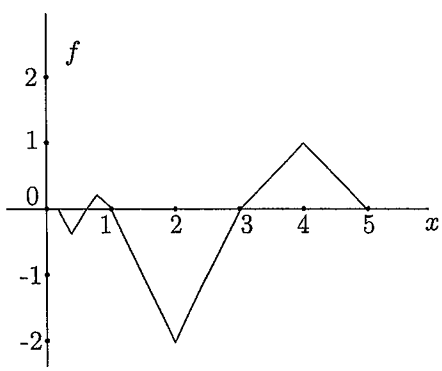
$$\varphi_i(x) := \int_0^{|x|} \chi_{A_i}(t)\,dt \quad \text{for } i = 1,2,3,4, \quad x \in \mathbb{R}.$$

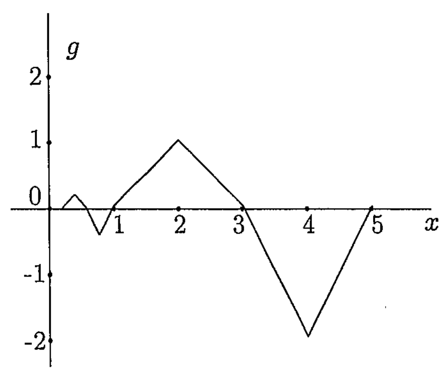
Finally, we define the cost and constraint functions in (6.1) by, respectively,

$$f(x) := -2\varphi_1(x) + 2\varphi_2(x) + \varphi_3(x) - \varphi_4(x) \quad \text{and} \quad g(x) := \varphi_1(x) - \varphi_2(x) - 2\varphi_3(x) + 2\varphi_4(x).$$

It is easy to see that the functions f and g are Lipschitz continuous on $\mathbb{R}$ with constant 2. Their graphs are given below. Observe that the point $x_0 = 0$ is an optimal solution to problem (6.1) Indeed, observe that $f(x) \geq 0$ for all feasible x and $f(x_0) = g(x_0) = 0$.

The figure below gives a part of the graphs of f and g. The complete graphs are obtained by applying an infinite number of homotheties with center at the origin of the part of the graph over the interval $[1,5]$, which is a piecewise affine function. We have $f(2\cdot 5^k) = g(4\cdot 5^k) = -2\cdot 5^k$ and $f(4\cdot 5^k) = g(2\cdot 5^k) = 5^k$ for all integer numbers k.





In this example conditions (6.2) and (6.4) reduce to

$$0 \in \partial_u\big(\lambda f + \mu g\big)(x_0) \;\text{ with }\; \lambda, \mu \geq 0 \;\text{ and }\; u = \pm 1. \tag{6.5}$$

where λ and μ are not zero simultaneously. Since the functions f and g are even, it is sufficient to consider only the case of $u = 1$ in (6.5).

Note that if there exist nonnegative multipliers λ and μ, not both zero, for which x_0 is an unconstrained local minimizer of the function $h := \lambda f + \mu g$, then (6.5) follows by Theorem 5.1.

However such multipliers do not exist. Indeed, assuming on the contrary that h has a local minimum at x_0 for some λ and μ of this type gives us, for all sufficiently small integer numbers ν, the following relationships:

$$0 = h(x_0) \leq h(2 \cdot 5^\nu) = \lambda f(2 \cdot 5^\nu) + \mu g(2 \cdot 5^\nu) = (-2\lambda + \mu)5^\nu$$
$$0 = h(x_0) \leq h(4 \cdot 5^\nu) = \lambda f(4 \cdot 5^\nu) + \mu g(4 \cdot 5^\nu) = (\lambda - 2\mu)5^\nu.$$

Adding the two inequalities obtained, we arrive at the contradiction $0 \leq -(\lambda + \mu)5^\nu < 0$.

As a consequence of the above inequality, we can choose either $y_\nu = 2 \cdot 5^\nu$ or $y_\nu = 4 \cdot 5^\nu$ to get $h(y_\nu) - h(x_0) \leq -\frac{1}{2}(\lambda + \mu)5^\nu$. To proceed further suppose, despite that x_0 is not a minimizer of h, that (6.5) holds with the choice $x_k = x_0$ like in the proof Theorem 5.1. Now for all number $\varepsilon > 0$ and some sequences $\varepsilon_k \to 0^+$ and $x_k^* \to 0$ as $k \to \infty$ we get

$$x_k^*(y_{\nu_k} - x_0) - \big(h(y_{\nu_k}) - h(x_0)\big) \leq (\varepsilon_k + \varepsilon)\big(|y_{\nu_k} - x_0| + |h(y_{\nu_k}) - h(x_0)|\big)$$

whenever the integer $\nu = \nu_k$ is sufficiently small. Recalling that $x_0 = 0$, the latter inequality can be rewritten in the form

$$x_k^* - \Big(1 - \varepsilon_k - \varepsilon\Big)\frac{h(y_{\nu_k})}{y_{\nu_k}} \leq \varepsilon_k + \varepsilon.$$

Passing now to the 'lim inf' as $k \to \infty$ with $0 < \varepsilon < 1$ and taking into account that $\frac{1}{2}(\lambda+\mu)\,5^{\nu_k} \leq -h(y_{\nu_k})$ and $0 < y_{\nu_k} \leq 4 \cdot 5^{\nu_k}$, we get $\frac{1}{8}(1-\varepsilon)(\lambda + \mu) \leq \varepsilon$. Passing further to the limit as $\varepsilon \to 0^+$ in the latter inequality leads us to the contradiction $\frac{1}{8}(\lambda + \mu) \leq 0$, which shows that the choice of $x_k = x_0$ in proving (6.5) does not work.

It remains to explore the case of $x_k \neq x_0$ to justify (6.5). Choosing, e.g., $x_k = 5^{-k}$ gives us

$$h(x) - h(x_k) = (-2\lambda + \mu)(x - x_k) \;\text{ whenever }\; x_k \leq x \leq x_k + 5^{-k}.$$

Consequently we get (6.5) if $-2\lambda + \mu \geq 0$. Similarly the directional subdifferential condition in (6.5) can be justified for $x_k = 2 \cdot 5^{-k}$ with $2\lambda - \mu \geq 0$, or for $x_k = 3 \cdot 5^{-k}$ with $\lambda - 2\mu \geq 0$, or for $x_k = 4 \cdot 5^{-k}$ with $-\lambda + 2\mu \geq 0$.

We learn a very instructive information from considering Example 6.1: to justify the directional necessary optimality conditions in (6.2) and (6.4), multipliers should be selected dependently on the choice of a sequence $x_k \neq x_0$. This idea leads us to the following major result of the section.

Theorem 6.2 (directional subdifferential necessary conditions in constrained optimization). *Let $x_0 \in \Omega \subset X$ be a local optimal solution to the constrained problem* (6.1), *where all the functions f and g_j as $j = 1, \ldots, p$ are locally Lipschitzian around x_0. Assume also that the set $\Omega \cap D_u(x_0, \Delta_0, \Delta_1)$ is of positive Lebesgue measure for all $0 \neq u \in T(x_0; \Omega)$. Then there are multipliers $\lambda \geq 0$ and $\mu_j \geq 0$ as $j = 1, \ldots, p$, not equal to zero simultaneously, satisfying conditions* (6.2) *and* (6.4) *in terms of the basic directional subdifferential* (3.2) *uniformly with respect to all the directions $0 \neq u \in X$.*

Proof. Some essential parts of the proof below hold under assumptions weaker that the local Lipschitz continuity of f and g_j around x_0 (like in Theorems 5.1 and Theorem 5.3); so we emphasize this in the proof. Denote by

$$I(x_0) := \big\{j \in \{1, \ldots, p\} \big|\, g_j(x_0) = 0\big\}$$

the collection of active indices at the reference point x_0 and select $\rho > 0$ so that $g_j(x) < 0$ for $j \notin I(x_0)$ and $x \in B(x_0, \rho)$ and that the set $\{(f(x), g_1(x), \ldots, g_p(x)) \mid x \in B(x_0, \rho)\}$ is bounded in $\mathbb{R} \times \mathbb{R}^p$. The existence of such a number ρ is ensured by the continuity of f and g_j at x_0. With no loss of generality, suppose that x_0 is a global minimizer of problem (6.1) with f and g_j restricted to $\Omega \cap B(x_0, \rho)$. In what follows we make use of the contingent cone (2.5) to the sets

$$T\big(-g_1(x_0), \ldots, g_p(x_0)); \mathbb{R}^p_+\big) = T\big(-g_1(x_0); \mathbb{R}_+\big) \times \ldots \times T\big(-g_p(x_0); \mathbb{R}_+\big)$$

with $T(-g_j(x_0); \mathbb{R}_+) = \mathbb{R}_+$ if $g_j(x_0) = 0$ and $T(-g_j(x_0); \mathbb{R}_+) = \mathbb{R}$ if $g_j(x_0) < 0$. Note that for $u \notin T(x_0; \Omega)$ the conclusion of the theorem is true with arbitrary multipliers λ and $\mu = (\mu_1, \ldots, \mu_p)$ since $\partial_u h(x_0) = X^*$. Hereafter we denote $h := \lambda f + \sum_{j=1}^p \mu_j g_j$.

Consider now the case of $u \in T(x_0; \Omega)$. We claim that *for all $\Delta_0, \Delta_1 > 0$ there is $x \in (\Omega \setminus \{x_0\}) \cap \overline{D}_u(x_0, \Delta_0, \Delta_1)$ with $\|x - x_0\| < \Delta_0$ such that whenever $\varepsilon > 0$ we can find $\delta(x) > 0$ depending also on ε with the properties that*

$$\overline{D}_u(x, \delta(x), \Delta_1) \subset \overline{D}_u(x_0, \Delta_0, \Delta_1)$$

and that for all $y \in \Omega \cap \overline{D}_u(x, \delta(x), \Delta_1)$ at least one of the following inequalities is satisfied:

$$f(y) - f(x) \geq -\varepsilon\|y - x\|, \quad g_j(y) - g_j(x) \geq -\varepsilon\|y - x\| \;\; \text{when} \;\; j \in I(x_0). \tag{6.6}$$

Assuming on the contrary that this claim does not hold for some $\varepsilon > 0$, we select $\Delta_0, \Delta_1 > 0$ such that $\Delta_0 \leq \rho$ and Δ_1 so small so that the relationships $x \in x_0 + C_{u+\Delta_1\overline{B}}$ and $y \in x + C_{u+\Delta_1\overline{B}}$ with $y \neq x$ imply that $\|y - x_0\| > \|x - x_0\|$. The existence of Δ_1 can be justified as in Theorem 5.3.

For $x \in (\Omega \setminus \{x_0\}) \cap \overline{D}_u(x_0, \Delta_0, \Delta_1)$ with $\|x - x_0\| < \Delta_0$ denote by $\mathcal{D}(x)$ the collections of all

$$y \in \Omega \cap (x + C_{u+\Delta_1\overline{B}}) \cap \overline{D}_u(x_0, \Delta_0, \Delta_1)$$

satisfying the conditions

$$f(y) - f(x) \leq -\varepsilon\|y - x\| \;\text{ and }\; g_j(y) - g_j(x) \leq -\varepsilon\|y - x\| \;\text{ whenever }\; j \in I(x_0). \tag{6.7}$$

It follows from the lower semicontinuity of the functions f and g_j that the set $\mathcal{D}(x) \neq \emptyset$ is closed; cf. the proof of Theorem 5.3. Thus the latter set is compact by $\dim X < \infty$. Since X is finite-dimensional, $\mathcal{D}(x)$ is also compact. Taking further $\overline{y} \in \mathcal{D}(x)$ such that $\|\overline{y} - x_0\| = \sup\{\|y - x_0\| \mid y \in \mathcal{D}(x)\}$, let us prove that $\|\overline{y} - x_0\| = \Delta_0$. If this is not the case, find $y \in \mathcal{D}(\overline{y})$ and observe that $y \in \mathcal{D}(x)$, which follows from (5.6), (5.7), and the fact that

$$g_j(y) - g_j(x_0) = \big(g_j(y) - g_j(\overline{y})\big) + \big(g_j(\overline{y}) - g_j(\overline{x})\big) \leq -\varepsilon\big(\|y - \overline{y}\| + \|\overline{y} - x\|\big) \leq -\varepsilon\|y - x\|$$

for all $j \in I(x_0)$. Remembering now the choice of Δ_1 yields that $\|y - x_0\| > \|\overline{y} - x_0\|$, a contradiction with the assumption imposed on $\overline{y}$.

Since $u \in T(x_0; \Omega)$, we choose $x_k \in (\Omega \setminus \{x_0\}) \cap \overline{D}_u(x_0, \Delta_0, \Delta_1)$ so that $x_k \overset{\Omega,u}{\to} x_0$ as $k \to \infty$ and then associate with x_k a vector $y_k \in \mathcal{D}(x_k)$ satisfying $\|y_k - x_0\| = \Delta_0$. Suppose without loss of generality that $y_k \to y_0$ as $k \to \infty$. Taking $\liminf_k$ in (6.7), we get that (5.8) holds and that

$$g_j(y_0) - g_j(x_0) \leq \liminf_k \big(g_j(y_k) - g_j(x_k)\big) \leq -\varepsilon\|y_0 - x_0\| = -\varepsilon\Delta_0 < 0$$

for all $j \in J(x_0)$, which contradicts the minimality of x_0 in (6.1) and thus justifiers the claim.

Observe that so far we did not use the Lipschitz property of f and g_j. Now we assume these functions are Lipschitz continuous on the whole space X. In fact it does not restrict the generality since the Kirszbraun theorem [8] ensures that any function Lipschitz continuous on the set $\Omega \cap \overline{B}(x_0, \rho)$ can be extended to the whole space as a Lipschitz function with the same constant. Furthermore, we know from the Rademacher theorem f and g_j are a.e. differentiable on X with respect to the Lebesgue measure. Thus we find a sequence $x_k \overset{\Omega\setminus\{x_0\},u}{\to} x_0$ such that f and g_j are differentiable at x_k for each $k \in \mathbb{N}$. Note that the sequences $\{f'(x_k)\}$ and $\{g_j'(x_k)\}$ are bounded, and thus we can assume that they converge as $k \to \infty$ to some vectors $a \in X$ and $b_j \in X$, respectively. It follows from (6.6) that for each k at least one of the following inequalities holds:

$$\langle f'(x_k), u\rangle \geq 0, \quad \langle g_j'(x_k), u\rangle \geq 0 \text{ for } j \in I(x_0),$$

which can be equivalently written in the form

$$\big(\langle f'(x_k), u\rangle, \langle g'(x_k), u\rangle\big) \notin -\text{int}\left(\mathbb{R}_+ \times T(-g(x_0); \mathbb{R}^p_+)\right) \text{ with } g := (g_1, \ldots, g_p).$$

Passing to the limit above as $k \to \infty$ gives us

$$\big(\langle a, u\rangle, \langle b, u\rangle\big) \notin -\text{int}\left(\mathbb{R}_+ \times T(-g(x_0); \mathbb{R}^p_+)\right). \tag{6.8}$$

Applying the classical separation theorem to (6.8), we find nonnegative numbers λ and μ_j as $j = 1, \ldots, p$, not all zero and independent of $u \in X$, such that $\mu_j g_j(x_0) = 0$ for $j = 1, \ldots, p$ and

$$\begin{gathered} \lambda\langle a, u\rangle + \textstyle\sum_{j=1}^p \mu_j \langle b_j, u\rangle \geq 0, \\ \lambda\alpha + \displaystyle\sum_{j=1}^p \mu_j \beta_j \geq 0 \text{ for all } (\alpha, \beta) \in \mathbb{R}_+ \times T\big(-g(x_0); \mathbb{R}^p_+\big), \end{gathered} \tag{6.9}$$

where $\beta := (\beta_1, \ldots, \beta_p)$. Thus we get multipliers $(\lambda, \mu_1, \ldots, \mu_p)$ satisfying (6.2), and it remains to show that they satisfy condition (6.4). To proceed, denote $c := \lambda a + \sum_{j=1}^p \mu_j b_j$ and observe that the first inequality in (6.9) can be written as $\langle c, u\rangle \geq 0$. Let $\varepsilon_k \to 0^+$ as $k \to \infty$ be such that $\|h'(x_k) - c\| \leq \frac{1}{2}\varepsilon_k$. Choose $\delta_k \to 0^+$ satisfying

$$\left\langle c, \frac{x - x_k}{\|x - x_k\|}\right\rangle \geq -\frac{1}{2}\varepsilon_k \text{ for } x - x_k \in C_{u+\delta_k\overline{B}}.$$

Finally, given $\varepsilon > 0$ and using the definition of $h'(x_k)$, we find $\delta(k) > 0$ for which

$$|h(x) - h(x_k) - \langle h'(x_k), x - x_k\rangle| \leq \varepsilon\|x - x_k\| \text{ whenever } \|x - x_k\| \leq \delta(k).$$

Combining the relationships above, we obtain

$$\begin{aligned} h(x) - h(x_k) &\geq \langle h'(x_k), x - x_k\rangle - \varepsilon\|x - x_k\| \\ &= \langle h'(x_k) - c, x - x_k\rangle + \langle c, x - x_k\rangle - \varepsilon\|x - x_k\| \\ &\geq -\|h'(x_k) - c\| \cdot \|x - x_k\| - \tfrac{1}{2}\varepsilon_k\|x - x_k\| - \varepsilon\|x - x_k\| \\ &\geq -(\varepsilon_k + \varepsilon)\|x - x_k\|. \end{aligned}$$

This implies with $x_k^* := 0$, $x^* := 0$, and $r \geq h(x)$ that

$$\begin{aligned} &\langle x_k^*, x - x_k\rangle - \big(r - h(x_k)\big) \leq -\big(h(x) - h(x_k)\big) \\ &\leq (\varepsilon_k + \varepsilon)\|x - x_k\| \leq \big(\varepsilon_k + \varepsilon\big)\big(\|x - x_k\| + |r - h(x_k)|\big), \end{aligned}$$

which ensures by (3.3) that $0 \in \partial_u h(x_0)$ and thus completes the proof of the theorem. $\square$

The following consequence of Theorem 6.2 and calculus rules of Section 4 presents a new result for classical constrained problems with smooth/strictly differentiable data.

Corollary 6.3 (necessary optimality conditions for constrained problems with strictly directionally differentiable data). *Let the functions f and g_j, $j = 1, \dots, p$, be strictly directionally differentiable at x_0 in the framework of Theorem 6.2. Then there are multipliers λ and μ_j as $j = 1, \dots, p$, not equal to zero simultaneously and independent of $u \in X$, such that conditions (6.2) and*

$$0 \in \lambda \partial_u f(x_0) + \sum_{j=1}^{p} \mu_j \partial_u g_j(x_0) \quad \text{for any} \quad 0 \neq u \in X \tag{6.10}$$

are satisfied, where the directional subdifferentials are computed by Theorem 4.3.

Proof. Since any function strictly differentiable at a point is locally Lipschitzian around this point, we derive condition (6.10) from that of (6.4) of Theorem 6.2 with the subsequent usage the calculus rules from Theorems 4.3 and Theorem 4.4. $\square$

Let us present a simple example showing that our new directional subdifferential necessary conditions allow us to recognize nonoptimal solutions while nondirectional subdifferential conditions fail.

Example 6.4 (comparison between directional and nondirectional subdifferential conditions). Consider a finite-dimensional constrained problem of type (6.1) given by

$$\text{minimize} \;\; f(x) \;\; \text{subject to} \;\; g(x) := -\|x\| \leq 0, \tag{6.11}$$

where $f\colon X \to \mathbb{R}$ is an arbitrary cost function that does not attain its local minimum at $x_0 = 0$. It is clear that the constrained problem (6.11) is equivalent to the unconstrained one of minimizing f on X, and so $x_0 = 0$ is not a local minimizer for (6.11). Assume for simplicity that f is strictly differentiable at x_0 and show that the nondirectional conditions (6.2) and (6.3) hold anyway at x_0 with some $(\lambda, \mu) \neq 0$, i.e., they are far removed from selecting nonoptimal points. Indeed, by the subdifferential sum rule from [11, Proposition 1.107] inclusion (6.3) is equivalent for the problem (6.11) under consideration to that of

$$0 \in \lambda f'(x_0) + \mu \partial(-\|x\|)(x_0) \;\; \text{with} \;\; \lambda, \mu \geq 0, \tag{6.12}$$

and the complementary slackness condition $\mu g(x_0) = 0$ is always satisfied at $x_0 = 0$. Furthermore, it follows from [11, Proposition 1.87 and Theorem 1.89] for the case of concave functions that

$$\partial(-\|x\|)(0) = S := \big\{x \in X \big|\; \|x\| = 1\big\}.$$

Substituting this into (6.12) gives us

$$0 \in \lambda f'(x_0) + \mu S \;\; \text{with} \;\; \lambda, \mu \geq 0, \tag{6.13}$$

which is obviously satisfied with $\lambda = 1$ and $\mu = \|f'(x_0)\|$ thus showing inefficiency of relationships (6.2) and (6.3) in terms the Mordukhovich limiting subdifferential (3.1) as necessary optimality conditions in problem (6.11). Since the latter subdifferential is known to be the smallest one among any nondirectional subdifferentials obeying natural requirements (see, e.g., [11] for precise results and discussions in this direction), other nondirectional subdifferentials also fail to produce efficient necessary conditions in the problem under consideration.

Now we apply the new directional subdifferential conditions from Theorem 6.3 to problem (6.11). Taking into account that every strictly differentiable function is directionally strictly differentiable at the corresponding point and applying calculus rules from Proposition 4.1 and Theorem 4.4 to the optimality condition (6.4), we get to $0 \in \lambda f'(x_0) + \mu \partial_u g(x_0)$ for some nonnegative multipliers λ and μ not both zero. Employing further Theorem 4.3 gives us relationships

$$0 = \lambda \langle f'(x_0), u \rangle + \mu \langle \xi, u \rangle \quad \text{for some} \quad \xi \in \partial_u g(x_0), \tag{6.14}$$

and thus $\langle \xi, u \rangle \le -\|u\|$ whenever $0 \neq u \in X$. It follows from (6.14) that $\lambda > 0$ and $\langle f'(x_0), u \rangle \ge 0$ for all nonzero u. Finally, from $\langle f'(x_0), u \rangle \ge 0$ and $\langle f'(x_0), -u \rangle \ge 0$ we obtain $\langle f'(x_0), u \rangle = 0$ for all $u \in X$, which reduces to the stationary condition $f'(x_0) = 0$. Thus is the right and expected first-order necessary optimality condition for the problem under consideration.

Note in conclusion that, similarly to the proof of Theorem 5.8 based on Lemma 5.7 and the arguments developed in the proof of Theorem 6.2, we can derive "no-gap" sufficient optimality conditions for the first-order isolated minimizers (5.11) in the constrained problem (6.1) in terms of the directional subdifferentials. Furthermore, directional subdifferential constructions seem to be useful in deriving second-order optimality conditions for the problems under consideration. This is a topic of our further research.

References

[1] A. Ben-Tal, J. Zowe: *Directional derivatives in nonsmooth optimization.* J. Optim. Theory Appl. 47 (1985), 483–490.

[2] R. W. Chaney: *Second-order necessary conditions in constrained semismooth optimization.* SIAM J. Control Optim. 25 (1987), 1072–1081.

[3] R. W. Chaney: *Second-order directional derivatives for nonsmooth functions.* J. Math. Anal. Appl. 128 (1987), 495–511.

[4] I. Ginchev: *Higher order optimality conditions in nonsmooth optimization.* Optimization 51 (2002), 47–72.

[5] I. Ginchev, A. Guerraggio: *Second-order optimality schemes for nonsmooth functions.* In: G. P. Crespi, A. Guerraggio, E. Miglierina, M. Rocca (Eds.), Recent Advances in Optimization, Proc. Varese, June 13–14, 2002, pp. 53–69, Datanova, Milano 2003.

[6] I. Ginchev, A. Guerraggio, M. Rocca: *First-order conditions for $C^{0,1}$ constrained vector optimization.* In: F. Giannessi, A. Maugeri (eds.), Variational analysis and applications, 427–450, Nonconvex Optim. Appl. 79, Springer, New York, 2005.

[7] I. Ginchev, B. Mordukhovich: *On directionally dependent subdifferentials.* C. R. Acad. Bulgare Sci. 64 (2011), 497–508.

[8] J. Heinonen: *Lectures on Lipschitz analysis.* Report, University of Jyväskylä Department of Mathematics and Statistics, 100. University of Jyväskylä, Jyväskylä, 2005.

[9] P. Q. Khanh, N. D. Tuan: *Optimality conditions for nonsmooth multiobjective optimization using Hadamard directional derivatives*, J. Optim. Theory Appl. 133 (2007), 341-357.

[10] B. S. Mordukhovich: *Maximum principle in problems of time optimal control with nonsmooth constraints.* J. Appl. Math. Mech. 40 (1976), 460–469.

[11] B. S. Mordukhovich: *Variational analysis and generalized differentiation. I. Basic theory.* Grundlehren der Mathematischen Wissenschaften, 330. Springer, Berlin, 2006.

[12] B. S. Mordukhovich: *Variational analysis and generalized differentiation. II. Applications.* Grundlehren der Mathematischen Wissenschaften, 331. Springer, Berlin, 2006.

[13] J. P. Penot: *The directional subdifferential of the difference of two convex functions.* J. Global Optim. 49 (2011), 505–519.

[14] R. T. Rockafellar: *Directional Lipschitzian functions and subdifferential calculus.* Proc. London Math. Soc. 39 (1079), 331-355.

[15] R. T. Rockafellar, R. J-B. Wets: *Variational analysis.* Springer, Berlin, 1998.